

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 20:17:35 ; Search time 42 Seconds  
(without alignments)  
907.187 Million cell updates/sec

Title: US-10-783-415-4  
Perfect score: 396  
Sequence: 1 MVDFAVDYKKNLYSDDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	396	100.0	396	2 I75615	mammary tumor inte
2	15	3.8	418	2 T45807	translation initia
3	8	2.0	136	2 PQ0517	pol protein - ovin
4	8	2.0	393	2 H81745	conserved hypothet
5	8	2.0	425	2 S32866	outL protein - Erw
6	8	2.0	432	2 T33118	hypothetical prote
7	8	2.0	436	2 B86469	protein F12K21.16
8	8	2.0	4092	1 S38128	dynamin heavy chain
9	7	1.8	65	2 B86846	hypothetical prote
10	7	1.8	98	2 B64097	hypothetical prote
11	7	1.8	123	1 T27492	cytochrome c 2C116
12	7	1.8	141	2 C24338	hemoglobin alpha-T
13	7	1.8	168	2 E81749	conserved hypothet
14	7	1.8	172	2 S56905	hypothetical prote
15	7	1.8	179	2 S05440	alpha-2u-globulin
16	7	1.8	181	1 UART	alpha-2u-globulin
17	7	1.8	181	2 I52504	alpha-2u-globulin
18	7	1.8	181	2 S05244	alpha-2u-globulin
19	7	1.8	181	2 I65319	alpha-2u-globulin
20	7	1.8	185	2 S75549	hypothetical prote
21	7	1.8	193	2 H83794	hypothetical prote
22	7	1.8	202	2 A83777	hypothetical prote
23	7	1.8	222	2 F70236	conserved hypothet
24	7	1.8	231	2 C69041	hypothetical prote
25	7	1.8	246	2 AG3644	flagellar biosynth
26	7	1.8	252	2 F90513	conserved hypothet
27	7	1.8	290	2 H83914	transcription regu
28	7	1.8	307	2 T01156	hypothetical prote
29	7	1.8	313	2 A34785	DNA-binding protei

30	7	1.8	332	2 JC5465	2,3-dihydroxybiphe
31	7	1.8	334	2 D85815	probable reductase
32	7	1.8	334	2 G64961	hypothetical prote
33	7	1.8	334	2 E90967	probable reductase
34	7	1.8	358	2 C42026	cyclic AMP respons
35	7	1.8	359	2 T34602	probable NTP prop
36	7	1.8	365	2 C69005	succinyl-CoA synth
37	7	1.8	367	2 T36116	probable oxidoredu
38	7	1.8	369	2 I78877	CAMP responsive el
39	7	1.8	375	2 JC8029	ornithine carbamoy
40	7	1.8	389	1 A39429	CAMP response elem
41	7	1.8	389	2 T47323	hypothetical prote
42	7	1.8	395	2 B71244	hypothetical prote
43	7	1.8	396	2 C75209	hypothetical prote
44	7	1.8	400	2 T33073	hypothetical prote
45	7	1.8	402	2 AD0148	tyrosine-specific

ALIGNMENTS

RESULT 1

I75615  
mammary tumor integration site 6 oncogene protein - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999  
C:Accession: I75615; I75614  
R:Marchetti, A.; Buttitta, F.; Miyazaki, S.; Gallahan, D.; Smith, G.H.; Callahan, R.  
J. Virol. 69, 1932-1938, 1995  
A:Title: Int-6, a highly conserved, widely expressed gene, is mutated by mouse mammary  
A:Reference number: I56914; MUID:95156630; PMID:7853537  
A:Accession: I75615  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-396 <MAR1>  
A:Cross-references: GB:S75224; NID:g913340; PID:g913341  
A:Accession: I75614  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 235-268, 'LKIKFPQSKGIRYVNOAVSGFMEIHRAAVRPEGHRLGRDPLEG' <MAR2>  
A:Cross-references: GB:S75223; NID:g913338; PIDN:AAC00047.1; PID:g913339  
C:Note: hypothetical mutant fusion protein  
C:Genetics:  
A:Gene: Int-6

Query Match	100.0%	Score 396;	DB 2;	Length 396;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 396;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MVDFAMDVYKKNLYSDDIPHALREKETTVAQIKQLOAETEPVKMFEDPETTRQMOSTRD	60	
DB	1	MVDFAMDVYKKNLYSDDIPHALREKETTVAQIKQLOAETEPVKMFEDPETTRQMOSTRD	60	
QY	61	GRMLFDYLADKHGFRQEQYLDLTRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSSL	120	
DB	61	GRMLFDYLADKHGFRQEQYLDLTRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSSL	120	
QY	121	WGKLAISEILMQNWDAAEMEDTLRLKETIDNNSSVSSPLQSLQORTLTIHWSLFFVFNHPKGR	180	
DB	121	WGKLAISEILMQNWDAAEMEDTLRLKETIDNNSSVSSPLQSLQORTLTIHWSLFFVFNHPKGR	180	
QY	181	DNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITKDKVRKROVLKDLVKVLCQESYTYK	240	
DB	181	DNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITKDKVRKROVLKDLVKVLCQESYTYK	240	
QY	241	DPITEFVCLYVNFDFDGAQKKLRECESVLVNDFELVACLEDFIENARLFIETFCRIHQ	300	
DB	241	DPITEFVCLYVNFDFDGAQKKLRECESVLVNDFELVACLEDFIENARLFIETFCRIHQ	300	
QY	301	CISINMLADKLMNTPPEAEERWIVNLIARNARLDAKIDSKLGHVVMGNNAVSPYQQVIEKTK	360	
DB	301	CISINMLADKLMNTPPEAEERWIVNLIARNARLDAKIDSKLGHVVMGNNAVSPYQQVIEKTK	360	

QY 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396  
|||||  
Db 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

## RESULT 2

T45807  
translation initiation factor 3-like protein - Arabidopsis thaliana  
N:Alternate names: protein F2809.140  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C:Accession: T45807  
R:Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z23014  
A:Accession: T45807  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-418 <BEN>  
A:Cross-references: UNIPROT:Q9M2L8; EMBL:AL1137080  
A:Experimental source: cultivar Columbia; BAC clone F2809  
C:Genetics:  
A:Map position: 3  
A:Introns: 74/1; 125/3; 201/3; 219/2; 292/3; 340/2; 386/2  
A:Note: F2809.140

Query Match 3.8%; Score 15; DB 2; Length 418;  
Best Local Similarity 100.0%; Pred. No. 4.1e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 LMGKLASBILMQNWD 134  
|||||  
Db 171 LMGKLASBILMQNWD 185

## RESULT 3

PQ0517  
pol protein - ovine lentivirus OLV-CV1 (strain CUI) (fragment)  
C:Species: ovine lentivirus OLV-CV1  
C>Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 09-Jul-2004  
C:Accession: PQ0517  
R:Campbell, B.J.; Thompson, D.R.; Williams, J.R.; Campbell, S.G.; Avery, R.J.  
J. Gen. Virol. 74, 201-210, 1993  
A:Title: Characterization of a New York ovine lentivirus isolate.  
A:Reference number: PQ0517; MUID:93155645; PMID:8381461  
A:Accession: PQ0517  
A:Molecule type: DNA  
A:Residues: 1-136 <CAM>  
A:Cross-references: UNIPROT:Q7LZ00  
C:Genetics:  
A:Gene: pol  
C:Superfamily: pol polyprotein  
C:Keywords: polyprotein

Query Match 2.0%; Score 8; DB 2; Length 136;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 VITNKDVR 219  
|||||  
Db 118 VITNKDVR 125

## RESULT 4

H81745  
conserved hypothetical protein TC0058 [imported] - Chlamydia muridarum (strain Nigg)  
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: H81745  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.P.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255; PMID:10684935  
A:Accession: H81745  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-393 <TET>  
A:Cross-references: UNIPROT:Q9PLP1; GB:AE002273; GB:AE002160; NID:g7190092; PIDN:AAF3894  
A:Experimental source: strain Nigg (MoPn)  
C:Genetics:  
A:Gene: TC0058

Query Match 2.0%; Score 8; DB 2; Length 393;  
Best Local Similarity 100.0%; Pred. No. 7.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 ALSSLWGK 123  
|||||  
Db 19 ALSSLWGK 26

## RESULT 5

S32866  
outL protein - Erwinia carotovora  
C:Species: Erwinia carotovora  
C>Date: 08-Dec-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C:Accession: S32866; S31755  
R:Reeves, P.J.; Whitcombe, D.; Wharham, S.; Gibson, M.; Allison, G.; Bunce, N.; Barallon,  
Mol. Microbiol. 8, 443-456, 1993  
A:Title: Molecular cloning and characterization of 13 out genes from Erwinia carotovora  
bacteria.  
A:Reference number: S32857; MUID:93316842; PMID:8326859  
A:Accession: S32866  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-425 <REE>  
A:Cross-references: UNIPROT:P31708; EMBL:X70049; NID:g42184; PIDN:CAA49653.1; PID:g42194  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, /./  
C:Genetics:  
A:Gene: outL  
C:Keywords: transmembrane protein

Query Match 2.0%; Score 8; DB 2; Length 425;  
Best Local Similarity 100.0%; Pred. No. 7.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RVLVPATD 113  
|||||  
Db 76 RVLVPATD 83

## RESULT 6

T33118  
hypothetical protein B0511.10 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T33118  
R:Tin-Wollam, A.; Sutterer, C.; Ozersky, P.  
submitted to the EMBL Data Library, May 1998  
A:Description: The sequence of C. elegans cosmid B0511.  
A:Reference number: Z21285  
A:Accession: T33118  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-432 <TIN>  
A:Cross-references: UNIPROT:O61820; EMBL:AF067608; PIDN:AAAC17647.1; GSPDB:GNO00019; CESP

A:Map position: 1  
A:Introns: 30/3; 122/3; 281/3

Query Match 2.0%; Score 8; DB 2; Length 432;  
Best Local Similarity 100.0%; Pred. No. 7.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 GKLAISEIL 129  
 Db 171 GKLAISEIL 178

RESULT 7  
 B86469  
 Protein F12K21.16 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: B86469  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: B86141; MUID:21016719; PMID:11130712  
 A:Accession: B86469  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-436 <STO>  
 A:Cross-references: UNIPROT:Q9LNL2; GB:AE005172; NID:98778261; PIDN:AAF79270.1; GSPDB:GN  
 C:Genetics:  
 A:Gene: F12K21.16  
 A:Map position: 1  
 C:Superfamily: Arabidopsis hypothetical protein F7N22.18

Query Match 2.0%; Score 8; DB 2; Length 436;  
 Best Local Similarity 100.0%; Pred. No. 8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 LPDYLDK 71  
 Db 372 LPDYLDK 379

RESULT 8  
 S38128  
 dynein heavy chain, cytosolic - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein YKR054c  
 N:Contains: dynein ATPase (EC 3.6.4.2)  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: S38128; S38130; S37701; S43936; S43077  
 R:Visers, S.; Urrestazu, L.A.; Jauniaux, J.C.  
 submitted to the Protein Sequence Database, March 1994  
 A:Reference number: S38118  
 A:Accession: S38128  
 A:Molecule type: DNA  
 A:Residues: 1-4092 <VIS>  
 A:Cross-references: UNIPROT:P36022; EMBL:Z28279; NID:9486510; PIDN:CAA82132.1; PID:94865  
 A:Experimental source: strain S288C  
 R:van Vliet-Reedijk, J.C.; Planta, R.J.  
 submitted to the Protein Sequence Database, March 1994  
 A:Reference number: S38130  
 A:Accession: S38130  
 A:Molecule type: DNA  
 A:Residues: 1-787 <VAN>  
 A:Cross-references: EMBL:Z28279; GSPDB:GN00011; MIPS:YKR054c  
 R:Li, Y.Y.; Yeh, E.Y.; Hays, T.; Bloom, K.S.  
 submitted to the EMBL Data Library, May 1993  
 A:Description: Disruption of mitotic spindle orientation in a yeast dynein mutant.  
 A:Reference number: S37701  
 A:Accession: S37701  
 A:Molecule type: DNA  
 A:Residues: 1-588 'C', 590-600, 'A', 602-1363, 'A', 1365-2631, 'P', 2633-2657, 2659, 'IGW', 2660-2  
 A:Cross-references: EMBL:U15626

R:Eshel, D.; Urrestazu, L.A.; Visers, S.; Jauniaux, J.C.; van Vliet-Reedijk, J.C.; P  
 Proc. Natl. Acad. Sci. U.S.A. 90, 11172-11176, 1993  
 A:Title: Cytoplasmic dynein is required for normal nuclear segregation in yeast.  
 A:Reference number: S43936; MUID:94068566; PMID:8248224  
 A:Accession: S43936  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-2117, 'IV', 2120-4092 <ESH>  
 A:Cross-references: EMBL:Z21877; NID:9439287; PIDN:CAA79923.1; PID:9439288  
 C:Genetics:  
 A:Gene: SGD:DYN1; DHCI; MIPS:YKR054c  
 A:Cross-references: SGD:S0001762; MIPS:YKR054c  
 A:Map position: 11R  
 C:Function:  
 A:Description: probably acts in cytoplasmic microtubule-based motile processes includin  
 red for nuclear segregation; required to achieve and maintain proper spindle positionin  
 C:Superfamily: dynein heavy chain, cytosolic  
 C:Keywords: ATP; heterotrimer; hydrolase; microtubule binding; nucleotide binding; P; R;  
 F:1796-1803/Region: nucleotide-binding motif A (P-loop)  
 F:2074-2081/Region: nucleotide-binding motif A (P-loop)  
 F:2418-2425/Region: nucleotide-binding motif A (P-loop)  
 F:2760-2767/Region: nucleotide-binding motif A (P-loop)  
 F:1802/Binding site: ATP (Lys) #status predicted  
 F:2080/Binding site: ATP (Lys) #status predicted  
 F:2424/Binding site: ATP (Lys) #status predicted  
 F:2766/Binding site: ATP (Lys) #status predicted

Query Match 2.0%; Score 8; DB 1; Length 4092;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 370 AMNIEKKL 377  
 Db 3594 AMNIEKKL 3601

RESULT 9  
 B86846  
 hypothetical protein yscD [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
 C:Species: Lactococcus lactis subsp. lactis  
 C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: B86846  
 R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrli  
 Genome Res. 11, 731-753, 2001  
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
 A:Reference number: A86625; MUID:21235186; PMID:11337471  
 A:Accession: B86846  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-65 <STO>  
 A:Cross-references: UNIPROT:Q9CER5; GB:AE005176; PID:gl2724793; PIDN:AAK05868.1; GSPDB:  
 A:Experimental source: strain IL1403  
 C:Genetics:  
 A:Gene: yscD

Query Match 1.8%; Score 7; DB 2; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 IIDLFLY 189  
 Db 51 IIDLFLY 57

RESULT 10  
 B64097  
 hypothetical protein HI0928 - Haemophilus influenzae (strain Rd KW20)  
 C:Species: Haemophilus influenzae  
 C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Jul-2004  
 C:Accession: B64097  
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, P  
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,  
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Snall, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: B64097  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-98 <TIGR>  
A:Cross-references: UNIPROT:P44887; GB:U32765; GB:L42023; NID:g1573838; PIDN:AAC22486.1;

Query Match 1.8%; Score 7; DB 2; Length 98;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LKQLOAE 38  
|||||  
DB 28 LKQLOAE 34

RESULT 11  
T27492  
Cytochrome c ZC116.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 03-Nov-2000 #sequence\_revision 03-Nov-2000 #text\_change 09-Jul-2004  
C:Accession: T27492  
R:Smye, R.  
submitted to the EMBL Data Library, June 1996  
A:Reference number: Z20376  
A:Accession: T27492  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-123 <WIL>  
A:Cross-references: UNIPROT:Q23240; EMBL:Z74046; PIDN:CAA98555.1; GSPDB:GN00023; CESP:ZC  
A:Experimental source: clone ZC116  
C:Genetics:  
A:Gene: CESP:ZC116.2  
A:Map position: 5  
A:Introns: 24/3; 94/3  
C:Superfamily: cytochrome c; cytochrome c homology  
C:Keywords: chromoprotein; heme; iron; metalloprotein  
F:20-113/Domain: cytochrome c homology <CYC>  
F:30.33/Binding site: heme (Cys) (covalent) #status predicted  
F:34.95/Binding site: heme iron (His, Met) (axial ligands) #status predicted

Query Match 1.8%; Score 7; DB 1; Length 123;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 LFDYLAD 70  
|||||  
DB 79 LFDYLAD 85

RESULT 12  
C24338  
hemoglobin alpha-T5 chain - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 08-Aug-1987 #sequence\_revision 08-Aug-1987 #text\_change 12-Jul-2004  
C:Accession: C24338  
R:Banville, D.; Williams, J.G.  
Nucleic Acids Res. 13, 5407-5421, 1985  
A:Title: The pattern of expression of the Xenopus laevis tadpole alpha-globin genes and  
A:Reference number: A93578; MUID:85297748; PMID:2993998  
A:Accession: C24338  
A:Molecule type: mRNA  
A:Residues: 1-141 <BAN>  
A:Cross-references: UNIPROT:P06638; GB:X02798; NID:G64542; PIDN:CAA26566.1; PID:G64543  
C:Comment: The initiator Met is not shown.  
C:Superfamily: globin; globin homology  
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxygen carrier  
F:2-141/Domain: Globin homology <GUB>  
F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 1.8%; Score 7; DB 2; Length 141;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 DDIPHAL 21  
|||||  
DB 74 DDIPHAL 80

RESULT 13  
E81749  
conserved hypothetical protein TC0041 [imported] - Chlamydia muridarum (strain Nigg)  
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: E81749  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150255; PMID:10684935  
A:Accession: E81749  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-168 <TET>  
A:Cross-references: UNIPROT:Q9PLQ4; GB:AE002271; GB:AE002160; NID:g7190063; PIDN:AAF3693  
A:Experimental source: strain Nigg (MoPn)  
C:Genetics:  
A:Gene: TC0041

Query Match 1.8%; Score 7; DB 2; Length 168;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 KLRCEES 268  
|||||  
DB 37 KLRCEES 43

RESULT 14  
S56905  
hypothetical protein YJL124c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein J0714  
C:Species: Saccharomyces cerevisiae  
C:Date: 05-May-1995 #sequence\_revision 08-Sep-1995 #text\_change 09-Jul-2004  
C:Accession: S56905  
R:Cziepluch, C.; Kordes, E.; Pujol, A.; Jauniaux, J.C.  
submitted to the Protein Sequence Database, September 1995  
A:Reference number: S56891  
A:Accession: S56905  
A:Molecule type: DNA  
A:Residues: 1-172 <CZI>  
A:Cross-references: UNIPROT:P47017; EMBL:Z49399; NID:g1008319; PIDN:CAA89419.1; PID:g100  
C:Genetics:  
A:Gene: SGD:LSM1; MIPS:YJL124c  
A:Cross-references: SGD:S0003660  
A:Map position: 10L  
C:Superfamily: hypothetical protein YJL124c

Query Match 1.8%; Score 7; DB 2; Length 172;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 RDGRMLF 65  
|||||  
DB 59 RDGRMLF 65

RESULT 15  
S05440  
alpha-2u-globulin precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 09-Jul-2004  
C:Accession: S05440; S70347

R;Gao, F.; Endo, H.; Yamamoto, M.  
 Nucleic Acids Res. 17, 4629-4636, 1989  
 A;Title: Length heterogeneity in rat salivary gland alpha-2-mu globulin mRNAs: multiple  
 A;Reference number: S05440; MUID:89315206; PMID:2473438  
 A;Accession: S05440  
 A;Status: translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-179 <GAO>  
 A;Cross-references: UNIPROT:063213; EMBL:X14552; NID:G55569; PIDN:CAA32690.1; PID:G55570  
 A;Experimental source: salivary gland  
 R;Bayard, C.; Holmquist, L.; Vesterberg, O.  
 Biochim. Biophys. Acta 1290, 129-134, 1996  
 A;Title: Purification and identification of allergenic alpha(2u)-globulin species of rat  
 A;Reference number: S70347; MUID:96254071; PMID:8645715  
 A;Accession: S70347  
 A;Molecule type: protein  
 A;Residues: 20-44 <BAY>  
 A;Experimental source: urine  
 C;Function:  
 A;Description: transport protein of hydrophobic molecules  
 A;Note: eight-stranded beta-barrel conformation; interior cavity lined with hydrophobic  
 C;Superfamily: lipocalin; lipocalin homology  
 C;Keywords: glycoprotein; lipid binding; salivary gland; urine  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-179/Product: alpha-2u-globulin #status experimental <MAT>  
 F;29-176/Domain: lipocalin homology <LIP>  
 F;54/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.8%; Score 7; DB 2; Length 179;  
 Best Local Similarity 100.0%; Pred.No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 RDNIIDL 186  
 |||||  
 Db 164 RDNIIDL 170

Search completed: March 8, 2005, 20:27:22  
 Job time : 44 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.


OM protein - protein search, using sw model

Run on: March 8, 2005, 20:08:14 ; Search time 165 Seconds  
(without alignments)  
928.225 Million cell updates/sec

Title: US-10-783-415-4

Perfect score: 396

Sequence: 1 MVDPMVYKLVYSDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table: 

Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	396	100.0	396	2 AAW02113	Aaw02113 Murine In
2	396	100.0	396	2 AAW02112	Aaw02112 Human hom
3	396	100.0	396	2 AAB47920	Aab47920 Murine In
4	396	100.0	396	2 AAB47921	Aab47921 Human Int
5	396	100.0	396	6 ABU05167	Abu05167 Human exp
6	396	100.0	396	6 ABU05169	Abu05169 Human exp
7	396	100.0	396	6 ABU05158	Abu05158 Human exp
8	396	100.0	396	7 ADJ70258	Adj70258 Human hea
9	396	100.0	396	8 ADN30996	Adn30996 Human Int
10	396	100.0	396	8 ADN30994	Adn30994 Murine In
11	396	100.0	445	5 ABB57306	Abb57306 Mouse isc
12	396	100.0	445	6 ABR33935	Abr33935 Human pro
13	396	100.0	445	6 ABU05160	Abu05160 Human exp
14	396	100.0	445	6 ABU05163	Abu05163 Human exp
15	396	100.0	445	6 ABU05157	Abu05157 Human exp
16	396	100.0	445	6 ABU05166	Abu05166 Human exp
17	396	100.0	445	6 ABU05159	Abu05159 Human exp
18	396	100.0	445	6 ABU05162	Abu05162 Human exp
19	396	100.0	445	6 ABU05164	Abu05164 Human exp
20	396	100.0	445	6 ABU05168	Abu05168 Human exp
21	396	100.0	445	7 ABM85381	Abm85381 Human pro
22	396	100.0	445	8 ADL13112	Adl13112 Human ste
23	396	100.0	445	8 ABM81005	Abm81005 Tumour-as
24	295	74.5	445	6 ABU05161	Abu05161 Human exp
25	204	51.5	320	7 ABM85380	Abm85380 Mouse pro

26	166	41.9	445	6	ABU05165	Abu05165 Human exp
27	103	26.0	117	5	ABP43106	Abp43106 Human ova
28	91	23.0	155	4	ABG21242	Abg21242 Novel hum
29	45	11.4	45	4	AAM22004	Aam22004 Peptide #
30	45	11.4	45	4	ABB44381	Abb44381 Peptide #
31	45	11.4	45	4	AAM38359	Aam38359 Peptide #
32	45	11.4	45	4	ABB27240	Abb27240 Protein #
33	45	11.4	45	4	AAM78125	Aam78125 Human bon
34	45	11.4	45	4	AAM65454	Aam65454 Human bra
35	45	11.4	45	4	ABG59761	Abg59761 Human liv
36	45	11.4	45	5	ABG47142	Abg47142 Human pep
37	43	10.9	1202	4	ABG21244	Abg21244 Novel hum
38	43	10.9	1202	7	ADF60389	Adf60389 Human con
39	42	10.6	42	4	AAM21570	Aam21570 Peptide #
40	42	10.6	42	4	ABB41600	Abb41600 Peptide #
41	42	10.6	42	4	ABB39233	Abb39233 Peptide #
42	42	10.6	42	4	AAM37849	Aam37849 Peptide #
43	42	10.6	42	4	AAM35393	Aam35393 Peptide #
44	42	10.6	42	4	ABB25431	Abb25431 Protein #
45	42	10.6	42	4	ABB26840	Abb26840 Protein #

## ALIGNMENTS

## RESULT 1

## AAW02113

ID AAW02113 standard; protein; 396 AA.

XX AAW02113;

DT 15-MAY-1997 (first entry)

DE Murine Int6 protein associated with MMTV integration and tumour growth.

KW MMTV; mouse mammary tumour virus; Int6; breast cancer; neoplasia;

KW diagnosis; treatment; immunotherapy; vaccine; probe; primer.

OS Mus musculus.

FH Key Location/Qualifiers

FT Misc-difference 27 /note= "potential cAMP/CGMP-dependent protein kinase phosphorylation site"

FT Misc-difference 51 /note= "potential protein kinase C phosphorylation site"

FT Misc-difference 57 /note= "potential casein kinase II phosphorylation site"

FT Misc-difference 90 /note= "potential tyrosine kinase phosphorylation site"

FT Misc-difference 94 /note= "potential glycosylation site"

FT Misc-difference 112 /note= "potential protein kinase C phosphorylation site"

FT Misc-difference 149 /note= "potential glycosylation site"

FT Misc-difference 214 /note= "potential casein kinase II phosphorylation site"

FT Misc-difference 238 /note= "potential casein kinase II phosphorylation site"

FT Misc-difference 314 /note= "potential casein kinase II phosphorylation site"

FT Misc-difference 363 /note= "potential protein kinase C phosphorylation site"

FT Misc-difference 381 /note= "potential casein kinase II phosphorylation site"

XX WO9624672-A1.

XX 15-AUG-1996.

XX 09-FEB-1996; 96WO-US001884.

```

PR 09-FEB-1995; 95US-00385998.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Marchetti A, Buttitta F, Smith GH, Callahan R;
XX
DR WPI; 1996-384444/38.
DR N-PSDB; AAT36177.
XX
XX DNA encoding Int6 tumour associated protein - and use of reagents derived
PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.
XX
PS Disclosure; Page 60-61; 93pp; English.
XX
XX AA02113 is the murine Int6 protein. The Int6 gene is located at
CC chromosome 15 of a mouse genome. The Int6 gene is associated with MMTV
CC (mouse mammary tumour virus) integration into a host genome during
CC tumorigenesis. Primers and probes may be derived from the Int6 gene
CC sequence and used for detection in assays to diagnose MMTV infection, or
CC any other Int6 gene integration. Antibodies against the Int6 protein can
CC be used in the same way. The DNA and protein may also be used gene
CC therapy and vaccines
XX
SQ Sequence 396 AA;
Query Match 100.0%; Score 396; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLQAETETPIVKMFEDPETTRQMSTRD 60
DB 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLQAETETPIVKMFEDPETTRQMSTRD 60
QY 61 GRMLFDYLDKKGFRQYDLTLRYAKFQYECGNYSAGAAEYLYFFRVLPATDRNALSSL 120
DB 61 GRMLFDYLDKKGFRQYDLTLRYAKFQYECGNYSAGAAEYLYFFRVLPATDRNALSSL 120
QY 121 WGLASEILMQNDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
DB 121 WGLASEILMQNDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQOYLNAIQTCMCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240
DB 181 DNIIDFLYQOYLNAIQTCMCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDPIENARLFIFETFCRIHQ 300
DB 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDPIENARLFIFETFCRIHQ 300
QY 301 CISINMLADKLNMTPEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
DB 301 CISINMLADKLNMTPEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
RESULT 2
AA02112
ID AA02112 standard; protein; 396 AA.
XX
AC AA02112;
XX
DT 14-MAY-1997 (first entry)
XX
DE Human homologue of Int6 protein associated with MMTV integration.
XX
KW MMTV; mouse mammary tumour virus; Int6; breast cancer; neoplasia;
KW diagnosis; treatment; immunotherapy; vaccine; probe; primer.
XX
OS Homo sapiens.
XX

PN WO9624672-A1.
XX
PD 15-AUG-1996.
XX
PF 09-FEB-1996; 96WO-US001884.
XX
PR 09-FEB-1995; 95US-00385998.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Marchetti A, Buttitta F, Smith GH, Callahan R;
XX
PI WPI; 1996-384444/38.
XX
DR N-PSDB; AAT36148.
XX
XX DNA encoding Int6 tumour associated protein - and use of reagents derived
PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.
XX
PS Claim 3; Page 62-63; 93pp; English.
XX
XX AA02112 represents the human homologue of murine Int6 protein. The Int6
CC gene is located at chromosome 15 of the mouse genome and is associated
CC with MMTV (mouse mammary tumour virus) integration into a host genome
CC during tumorigenesis. Primers and probes may be derived from the Int6
CC cDNA sequence and used for the detection of the Int6 gene. These can be
CC used in assays to diagnose MMTV infection, or any other Int6 gene
CC integration. Antibodies against the Int6 protein can be used in the same
CC way
XX
SQ Sequence 396 AA;
Query Match 100.0%; Score 396; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLQAETETPIVKMFEDPETTRQMSTRD 60
DB 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLQAETETPIVKMFEDPETTRQMSTRD 60
QY 61 GRMLFDYLDKKGFRQYDLTLRYAKFQYECGNYSAGAAEYLYFFRVLPATDRNALSSL 120
DB 61 GRMLFDYLDKKGFRQYDLTLRYAKFQYECGNYSAGAAEYLYFFRVLPATDRNALSSL 120
QY 121 WGLASEILMQNDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
DB 121 WGLASEILMQNDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQOYLNAIQTCMCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240
DB 181 DNIIDFLYQOYLNAIQTCMCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDPIENARLFIFETFCRIHQ 300
DB 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDPIENARLFIFETFCRIHQ 300
QY 301 CISINMLADKLNMTPEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
DB 301 CISINMLADKLNMTPEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
RESULT 3
AA047920
ID AA047920 standard; protein; 396 AA.
XX
AC AA047920;
XX
DT 21-MAY-2002 (first entry)
XX
DE Murine Int6.

```



XX Murine; human; Int6; integration site; deregulation; neoplasia;  
 KW mouse mammary tumour virus; MMV; cancer; immunotherapy; gene therapy;  
 KW prenatal screening; foetus; vaccine.  
 XX Mus musculus.

XX Location/Qualifiers  
 XX 27 /note= "Phosphorylation site for cAMP/cGMP-dependent  
 FT protein kinase"  
 FT 51  
 FT Modified-site  
 FT 57 /note= "Phosphorylation site for protein kinase C"  
 FT Modified-site  
 FT 90 /note= "Phosphorylation site for casein kinase II"  
 FT Modified-site  
 FT 94 /note= "Phosphorylation site for tyrosine kinase"  
 FT Modified-site  
 FT /note= "Phosphorylation site for tyrosine kinase and  
 FT glycosylation site"  
 FT 112 /note= "Phosphorylation site for protein kinase C"  
 FT Modified-site  
 FT 149 /note= "Phosphorylation site for tyrosine kinase and  
 FT glycosylation site"  
 FT 214  
 FT Modified-site  
 FT /note= "Phosphorylation site for casein kinase II"  
 FT 238  
 FT Modified-site  
 FT /note= "Phosphorylation site for casein kinase II"  
 FT 314  
 FT Modified-site  
 FT /note= "Phosphorylation site for casein kinase II"  
 FT 363  
 FT Modified-site  
 FT /note= "Phosphorylation site for protein kinase C"  
 FT 381  
 FT Modified-site  
 FT /note= "Phosphorylation site for casein kinase II"  
 FT XX  
 XX US6342392-B1.  
 XX 29-JAN-2002.  
 XX 23-AUG-1999; 99US-00378842.  
 XX 09-FEB-1995; 95US-00385998.  
 XX 09-FEB-1996; 96WO-US001884.  
 XX 25-SEP-1997; 97US-00875847.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX Marchetti A, Buttitta F, Smith GH, Callahan R;  
 XX WPI; 1996-384444/38.  
 XX N-PSDB; AAI72499.  
 XX DNA encoding Int6 tumour associated protein - and use of reagents derived  
 PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.  
 XX Disclosure; Col 35-38; 45pp; English.

XX This sequence shows murine Int6. Int6 is an integration site for mouse  
 CC mammary tumour virus (MMTV), which causes deregulation of expression of  
 CC cellular genes adjacent to the site of MMTV integration in mammary  
 CC tumours. The Int6 protein has been found to be highly conserved across  
 CC species, with Drosophila Int6 being 60% identical to human/mouse Int6.  
 CC This indicates that Int6 is serving a basic life function. The method of  
 CC the invention comprises assaying a sample to detect a human Int6 nucleic  
 CC acid sequence, or its fragment, by contacting the sample with a sequence  
 CC of at least 15 consecutive nucleotides of human Int6 cDNA or a  
 CC conservative variant of it, where a disrupted expression or loss of  
 CC expression of the variant is associated with neoplasia. The method is  
 CC useful for prenatal screening of a foetus or to pre-symptomatically  
 CC screen a subject at risk of having cancer. Detecting mutations in the  
 CC Int6 gene can provide diagnostic and prognostic information. The nucleic  
 CC acids and proteins are useful in immunotherapy, gene therapy or as

CC vaccines for treating or preventing cancer. The nucleic acids are useful  
 CC as probes for isolating homologues of Int6 gene or for detecting  
 CC mutations in the Int6 gene  
 XX  
 SQ Sequence 396 AA;  
 Query Match 100.0%; Score 396; DB 2; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVAQLKQLAETEPVVKMFEDPETTRQMQSTRD 60  
 DB 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVAQLKQLAETEPVVKMFEDPETTRQMQSTRD 60  
 QY 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVPAATDRNALSSL 120  
 DB 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVPAATDRNALSSL 120  
 QY 121 WGLASEILLQNDAAAMEDLTRKETIDNNSVSPLOQRTWLIIHWSLVFFFNHPKGR 180  
 DB 121 WGLASEILLQNDAAAMEDLTRKETIDNNSVSPLOQRTWLIIHWSLVFFFNHPKGR 180  
 QY 181 DNIIDFLYQPYQLNAIQTMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTK 240  
 DB 181 DNIIDFLYQPYQLNAIQTMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTK 240  
 QY 241 DPITEFVECLYVNFDFDGAOKLRECESVLVNDFFLVACLEDFTIENARLFIPTFCRIHQ 300  
 DB 241 DPITEFVECLYVNFDFDGAOKLRECESVLVNDFFLVACLEDFTIENARLFIPTFCRIHQ 300  
 QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVGNNNAVSPYQOVIEKTK 360  
 DB 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVGNNNAVSPYQOVIEKTK 360  
 QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396  
 DB 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 4  
 AAB47921  
 ID AAB47921 standard; protein; 396 AA.  
 AC AAB47921;  
 XX 21-MAY-2002 (first entry)  
 XX Human Int6.  
 XX Murine; human; Int6; integration site; deregulation; neoplasia;  
 KW mouse mammary tumour virus; MMTV; cancer; immunotherapy; gene therapy;  
 KW prenatal screening; foetus; vaccine; chromosome 8q22-q24.  
 XX Homo sapiens.  
 XX US6342392-B1.  
 XX 29-JAN-2002.  
 XX 23-AUG-1999; 99US-00378842.  
 XX 09-FEB-1995; 95US-00385998.  
 XX 09-FEB-1996; 96WO-US001884.  
 XX 25-SEP-1997; 97US-00875847.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX Marchetti A, Buttitta F, Smith GH, Callahan R;  
 XX WPI; 1996-384444/38.  
 XX N-PSDB; AAI72499.  
 XX DNA encoding Int6 tumour associated protein - and use of reagents derived  
 PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.

PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.  
PS Disclosure; Col 39-44; 45pp; English.  
XX  
XX This sequence shows human Int6. The human Int6 coding sequence was  
CC isolated by using murine Int6 sequences as probes. Human Int6 is  
CC organised into 13 exons as is murine Int6, and contains a CA-repeat in  
CC the 7th intron. Human Int6 has been localised to chromosome 8, more  
CC specifically to 8q22-q24. Int6 is an integration site for mouse mammary  
CC tumour virus (MMTV), which causes deregulation of expression of cellular  
CC genes adjacent to the site of MMTV integration in mammary tumours. The  
CC Int6 protein has been found to be highly conserved across species, with  
CC Drosophila Int6 being 60% identical to human/mouse Int6. This indicates  
CC that Int6 is serving a basic life function. The method of the invention  
CC comprises assaying a sample to detect a human Int6 nucleic acid sequence,  
CC or its fragment, by contacting the sample with a sequence of at least 15  
CC consecutive nucleotides of human Int6 cDNA or a conservative variant of  
CC it, where a disrupted expression or loss of expression of the variant is  
CC associated with neoplasia. The method is useful for prenatal screening of  
CC a foetus or to pre-symptomatically screen a subject at risk of having  
CC cancer. Detecting mutations in the Int6 gene can provide diagnostic and  
CC prognostic information. The nucleic acids and proteins are useful in  
CC immunotherapy, gene therapy or as vaccines for treating or preventing  
CC cancer. The nucleic acids are useful as probes for isolating homologues  
CC of Int6 gene or for detecting mutations in the Int6 gene  
XX  
SQ Sequence 396 AA;  
Query Match 100.0%; Score 396; DB 2; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVDFAVDVKNLYSDDI PHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMSTRD 60  
Db 1 MVDFAVDVKNLYSDDI PHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMSTRD 60  
QY 61 GRMLFDYLDKKGFRQYEDTLRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120  
Db 61 GRMLFDYLDKKGFRQYEDTLRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120  
QY 121 WGLASEILMOWNDAAEMEDLTRKETIDNNSVSSPLOSQOORTWLHWSLFFVFNHPKGR 180  
Db 121 WGLASEILMOWNDAAEMEDLTRKETIDNNSVSSPLOSQOORTWLHWSLFFVFNHPKGR 180  
QY 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDKVRKROVLKDLVKVIOQESYTYK 240  
Db 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDKVRKROVLKDLVKVIOQESYTYK 240  
QY 241 DPITEFVECLYVNFDPDGAQKKLRCECVLVNDFFLVACLEDPIENARLFIPTFCRIHQ 300  
Db 241 DPITEFVECLYVNFDPDGAQKKLRCECVLVNDFFLVACLEDPIENARLFIPTFCRIHQ 300  
QY 301 CISINMLADKLNMTPEAEERWVNLIRNARLDKIDSKLGHVVMGNVSPVQOVIETKY 360  
Db 301 CISINMLADKLNMTPEAEERWVNLIRNARLDKIDSKLGHVVMGNVSPVQOVIETKY 360  
QY 361 SLFSRQMLAMNIEKKLNQSRSEAPNWTQDSGFY 396  
Db 361 SLFSRQMLAMNIEKKLNQSRSEAPNWTQDSGFY 396  
RESULT 5  
ID ABU05167  
XX ABU05167 standard; protein; 396 AA.  
AC ABU05167;  
XX  
XX 29-JAN-2003 (first entry)  
DE Human expressed protein tag (EPT) #1833.  
XX  
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;

KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
XX Homo sapiens.  
XX  
XX WO200278524-A2.  
XX  
XX 10-OCT-2002.  
XX  
XX 28-MAR-2002; 2002WO-US009671.  
XX  
XX 28-MAR-2001; 2001US-0279495P.  
XX  
XX 21-MAY-2001; 2001US-0292544P.  
XX  
XX 08-AUG-2001; 2001US-0310801P.  
XX  
XX 01-OCT-2001; 2001US-0326370P.  
XX  
XX 04-DEC-2001; 2001US-0336780P.  
XX  
XX 20-FEB-2002; 2002US-0358985P.  
XX  
XX (ZYCO-) ZYCOS INC.  
XX  
XX Chicx RM, Tomlinson AJ, Urban RG;  
XX  
XX WPI; 2003-040607/03.  
XX  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.  
XX  
XX Example 2; SEQ ID NO 1833; 134pp; English.  
XX  
XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX transporter, cytoskeletal protein, receptor or transcription factor. The  
XX polypeptide is useful as an immunogenic composition for eliciting in a  
XX mammal an immunogenic response directed against any of the purified  
XX polypeptide. The purified polypeptide, or the antibody that binds to this  
XX polypeptide, is useful for treating cancer. The polypeptide is also  
XX useful for identifying compounds that binds to a naturally processed  
XX class I or class II MHC-binding polypeptide. The polypeptides and  
XX polynucleotides are particularly useful for treating or preventing  
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
XX lymphoma or leukaemia. These are also useful for screening agents for  
XX treating the above mentioned diseases. This sequence represents an  
XX expressed protein tag (EPT) isolated from human tissue for translational  
XX profiling. Note: This sequence does not appear in the printed  
XX specification but was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 396 AA;  
Query Match 100.0%; Score 396; DB 6; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVDFAVDVKNLYSDDI PHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMSTRD 60  
Db 1 MVDFAVDVKNLYSDDI PHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMSTRD 60  
QY 61 GRMLFDYLDKKGFRQYEDTLRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120  
Db 61 GRMLFDYLDKKGFRQYEDTLRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120  
QY 121 WGLASEILMOWNDAAEMEDLTRKETIDNNSVSSPLOSQOORTWLHWSLFFVFNHPKGR 180  
Db 121 WGLASEILMOWNDAAEMEDLTRKETIDNNSVSSPLOSQOORTWLHWSLFFVFNHPKGR 180  
QY 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDKVRKROVLKDLVKVIOQESYTYK 240  
Db 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDKVRKROVLKDLVKVIOQESYTYK 240  
QY 241 DPITEFVECLYVNFDPDGAQKKLRCECVLVNDFFLVACLEDPIENARLFIPTFCRIHQ 300

```
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVGNNNAVSPYQVIEKTK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVGNNNAVSPYQVIEKTK 360
Qy 361 SLSFRSOMLANNIEKKLNQNSRSEAPNATQDSGFY 396
Db 361 SLSFRSOMLANNIEKKLNQNSRSEAPNATQDSGFY 396

RESULT 6
ABU05169
ID ABU05169 standard; protein; 396 AA.
XX AC ABU05169;
XX XX
DT 29-JAN-2003 (first entry)
XX DE Human expressed protein tag (EPT) #1835.
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX OS Homo sapiens.
XX PN WO200278524-A2.
XX PD 10-OCT-2002.
XX PF 28-MAR-2002; 2002WO-US009671.
XX PR 28-MAR-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.
XX PR 01-OCT-2001; 2001US-0326370P.
XX PR 04-DEC-2001; 2001US-0336780P.
XX PR 20-FEB-2002; 2002US-0358985P.
XX PA (ZYCO-) ZYCOS INC.
XX XX
XX PI Chicx RM, Tomlinson AJ, Urban RG;
XX XX
XX WPI; 2003-040607/03.
XX XX
XX PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX XX
XX PS Example 2; SEQ ID NO 1835; 134pp; English.
XX CC
XX CC The invention describes a purified polypeptide, which comprises a
XX CC fragment of a kinase, phosphatase, protease, protease inhibitor,
XX CC transporter, cytoskeletal protein, receptor or transcription factor. The
XX CC polypeptide is useful as an immunogenic composition for eliciting in a
XX CC mammal an immunogenic response directed against any of the purified
XX CC polypeptide. The purified polypeptide, or the antibody that binds to this
XX CC polypeptide, is useful for treating cancer. The polypeptide is also
XX CC useful for identifying compounds that binds to a naturally processed
XX CC class I or class II MHC-binding polypeptide. The polypeptides and
XX CC polynucleotides are particularly useful for treating or preventing
XX CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX CC lymphoma or leukaemia. These are also useful for screening agents for
XX CC treating the above mentioned diseases. This sequence represents an
XX CC expressed protein tag (EPT) isolated from human tissue for translational
XX CC profiling. Note: This sequence does not appear in the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
```

```
XX SQ Sequence 396 AA;
Query Match 100.0%; Score 396; DB 6; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVDFAVDVYKLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMQSTRD 60
Db 1 MVDFAVDVYKLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMQSTRD 60
Qy 61 GRMLFDYLADKHGFRQYELDTLYRYAKFQYECGNYSGAAEYLYFRVLVPATDRNALSSL 120
Db 61 GRMLFDYLADKHGFRQYELDTLYRYAKFQYECGNYSGAAEYLYFRVLVPATDRNALSSL 120
Qy 121 WGKLASEILMNWDAAWEDLTRKETIDNNSVSPLOSQORTWLIIHWSLFPVFNHPKGR 180
Db 121 WGKLASEILMNWDAAWEDLTRKETIDNNSVSPLOSQORTWLIIHWSLFPVFNHPKGR 180
Qy 181 DNIIDFLYQOYVNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTK 240
Db 181 DNIIDFLYQOYVNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVGNNNAVSPYQVIEKTK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVGNNNAVSPYQVIEKTK 360
Qy 361 SLSFRSOMLANNIEKKLNQNSRSEAPNATQDSGFY 396
Db 361 SLSFRSOMLANNIEKKLNQNSRSEAPNATQDSGFY 396

RESULT 7
ABU05158
ID ABU05158 standard; protein; 396 AA.
XX AC ABU05158;
XX XX
DT 29-JAN-2003 (first entry)
XX XX
XX DE Human expressed protein tag (EPT) #1824.
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX KW protease inhibitor; transporter; cytoskeletal protein;
XX KW receptor; transcription factor; cancer; MHC;
XX KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX OS Homo sapiens.
XX PN WO200278524-A2.
XX PD 10-OCT-2002.
XX PF 28-MAR-2002; 2002WO-US009671.
XX PR 28-MAR-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.
XX PR 01-OCT-2001; 2001US-0326370P.
XX PR 04-DEC-2001; 2001US-0336780P.
XX PR 20-FEB-2002; 2002US-0358985P.
XX XX
XX XX (ZYCO-) ZYCOS INC.
XX XX
XX XX Chicx RM, Tomlinson AJ, Urban RG;
XX XX
XX WPI; 2003-040607/03.
XX XX
```

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX  
PS Example 2; SEQ ID NO 1824; 134pp; English.  
XX  
CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 396 AA;

Query Match 100.0%; Score 396; DB 6; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPTTQMOSTRD 60  
DB 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPTTQMOSTRD 60  
QY 61 GRMLFDYLDKKGFRQYLDLTVYAKFQYECGNYSGAAEYLYFFRVLPATRNALSSL 120  
DB 61 GRMLFDYLDKKGFRQYLDLTVYAKFQYECGNYSGAAEYLYFFRVLPATRNALSSL 120  
QY 121 WGLKASEILMNQNDAAEMDLTRKETIDNNSVSPQLSQLOORTWLHWSLFFVFNHPKGR 180  
DB 121 WGLKASEILMNQNDAAEMDLTRKETIDNNSVSPQLSQLOORTWLHWSLFFVFNHPKGR 180  
QY 181 DNIIDLFLYQPYQLNAIQWCPHILRYLTAVITNKDVKRRQVLKDLVKVIOQESYTYK 240  
DB 181 DNIIDLFLYQPYQLNAIQWCPHILRYLTAVITNKDVKRRQVLKDLVKVIOQESYTYK 240  
QY 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300  
DB 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300  
QY 301 CISINMLADKLNTPBEAERWIVNLIRNARLDKIDSKLGHVYVGNNAVSPYQOVIEKTK 360  
DB 301 CISINMLADKLNTPBEAERWIVNLIRNARLDKIDSKLGHVYVGNNAVSPYQOVIEKTK 360  
QY 361 SLFSFRSOMLAWNTEKKLNQNSRSEAPNWTQDSGFY 396  
DB 361 SLFSFRSOMLAWNTEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 8  
ADJ70258  
ID ADJ70258 standard; protein; 396 AA.  
XX  
AC ADJ70258;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Human heat mitochondrial protein as a therapeutic target SeqID2064.  
XX  
DE  
XX  
KW mitochondrial; human; screening assay; diabetes mellitus;  
KW Huntington's disease; osteoarthritis; LHON;  
KW Leber's hereditary optic neuropathy; LHON;

KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;  
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;  
KW osteopathic; ophthalmological; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO2003087768-A2.  
XX  
PD 23-OCT-2003.  
XX  
PF 04-APR-2003; 2003WO-US010870.  
XX  
PR 12-APR-2002; 2002US-0372843P.  
PR 17-JUN-2002; 2002US-0389987P.  
PR 20-SEP-2002; 2002US-0412418P.  
XX  
PA (MITO-) MITOKOR.  
PA (BUCK-) BUCK INST AGE RES.  
XX  
PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;  
PI Warnock DE;  
XX  
DR WPI; 2003-845369/78.  
XX  
PT Identifying a mitochondrial target for drug screening assays and for  
PT treating diseases associated with altered mitochondrial function,  
PT comprises detecting a modified polypeptide in a sample and correlating  
PT with the disease.  
XX  
PS Claim 1; SEQ ID NO 2064; 180pp; English.  
XX  
CC This invention relates to novel mitochondrial targets that can be used  
CC for therapeutic intervention in treating a disease associated with  
CC altered mitochondrial function. Specifically, it refers to a method for  
CC identifying proteins of the human heart mitochondrial proteome that are  
CC useful for drug screening assays, as well as therapeutic targets. The  
CC present invention describes a method for identifying such proteins that  
CC can be used in the treatment of various diseases associated with altered  
CC mitochondrial function including diabetes mellitus, Huntington's disease,  
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
CC compositions have neuroprotective, nontropic, antidiabetic,  
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
CC cytostatic activities. This polypeptide sequence is a human heart  
CC mitochondrial protein of the invention.  
XX  
SQ Sequence 396 AA;

Query Match 100.0%; Score 396; DB 7; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPTTQMOSTRD 60  
DB 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPTTQMOSTRD 60  
QY 61 GRMLFDYLDKKGFRQYLDLTVYAKFQYECGNYSGAAEYLYFFRVLPATRNALSSL 120  
DB 61 GRMLFDYLDKKGFRQYLDLTVYAKFQYECGNYSGAAEYLYFFRVLPATRNALSSL 120  
QY 121 WGLKASEILMNQNDAAEMDLTRKETIDNNSVSPQLSQLOORTWLHWSLFFVFNHPKGR 180  
DB 121 WGLKASEILMNQNDAAEMDLTRKETIDNNSVSPQLSQLOORTWLHWSLFFVFNHPKGR 180  
QY 181 DNIIDLFLYQPYQLNAIQWCPHILRYLTAVITNKDVKRRQVLKDLVKVIOQESYTYK 240  
DB 181 DNIIDLFLYQPYQLNAIQWCPHILRYLTAVITNKDVKRRQVLKDLVKVIOQESYTYK 240  
QY 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300  
DB 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300

QY 301 CINSINMLADKLNMTPEAEERWIVNLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
 DB 301 CINSINMLADKLNMTPEAEERWIVNLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
 QY 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396  
 DB 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 9  
 ID ADN30996  
 AC ADN30996  
 DT 29-JUL-2004 (first entry)  
 DE Human Int6 protein.  
 KW Human; Int6; mammary epithelial cellular growth; cancer; cytostatic.  
 OS Homo sapiens.  
 PN US6737251-B2.  
 PD 18-MAY-2004.  
 PF 14-MAY-2001; 2001US-00858152.  
 PR 09-FEB-1995; 95US-00385998.  
 PR 09-FEB-1996; 96US-00875847.  
 PR 09-FEB-1996; 96WO-US001884.  
 PR 23-AUG-1999; 99US-00378842.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX Marchetti A, Buttitta F, Smith GH, Callahan R;  
 DR WPI; 2004-387097/36.  
 DR N-PSDB; ADN30995.  
 XX Novel tumor Int6 recombinant protein that deregulates mammary epithelial  
 PT cellular growth, useful for treating cancer.  
 XX Claim 1; SEQ ID NO 4; 44pp; English.  
 XX The invention relates to the Int6 protein and the cDNA encoding it. The  
 CC Int6 protein deregulates mammary epithelial cellular growth. The cDNA and  
 CC protein are useful as vaccines for treating cancer. This sequence  
 CC represents the human Int6 protein of the invention.  
 XX  
 SQ Sequence 396 AA;  
 Query Match 100.0%; Score 396; DB 8; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLQAEETPIVVMFEDPETTRMQSTRD 60  
 DB 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLQAEETPIVVMFEDPETTRMQSTRD 60  
 QY 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVFPATDNLSSL 120  
 DB 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVFPATDNLSSL 120  
 QY 121 WGLKASILMKNQNDAAWEDLTRKETIDNNSVSPLOSQOORTWLIHWSLFPVFNHPKGR 180  
 DB 121 WGLKASILMKNQNDAAWEDLTRKETIDNNSVSPLOSQOORTWLIHWSLFPVFNHPKGR 180  
 QY 181 DNIIDLFLYQPOYLNAIQTCWPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240  
 DB 181 DNIIDLFLYQPOYLNAIQTCWPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240

QY 241 DPITERVECLYVNFDFDGAQKLCRECSVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300  
 DB 241 DPITERVECLYVNFDFDGAQKLCRECSVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300  
 QY 301 CINSINMLADKLNMTPEAEERWIVNLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
 DB 301 CINSINMLADKLNMTPEAEERWIVNLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
 QY 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396  
 DB 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 10  
 ID ADN30994  
 AC ADN30994  
 DT 29-JUL-2004 (first entry)  
 DE Murine Int6 protein.  
 KW Mouse; Int6; mammary epithelial cellular growth; cancer; cytostatic.  
 OS Mus musculus.  
 PN US6737251-B2.  
 PD 18-MAY-2004.  
 PF 14-MAY-2001; 2001US-00858152.  
 PR 09-FEB-1995; 95US-00385998.  
 PR 09-FEB-1996; 96US-00875847.  
 PR 09-FEB-1996; 96WO-US001884.  
 PR 23-AUG-1999; 99US-00378842.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX Marchetti A, Buttitta F, Smith GH, Callahan R;  
 DR WPI; 2004-387097/36.  
 DR N-PSDB; ADN30993.  
 XX Novel tumor Int6 recombinant protein that deregulates mammary epithelial  
 PT cellular growth, useful for treating cancer.  
 XX Example 4; SEQ ID NO 2; 44pp; English.  
 XX The invention relates to the Int6 protein and the cDNA encoding it. The  
 CC Int6 protein deregulates mammary epithelial cellular growth. The cDNA and  
 CC protein are useful as vaccines for treating cancer. This sequence  
 CC represents the murine Int6 protein of the invention.  
 XX  
 SQ Sequence 396 AA;  
 Query Match 100.0%; Score 396; DB 8; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLQAEETPIVVMFEDPETTRMQSTRD 60  
 DB 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLQAEETPIVVMFEDPETTRMQSTRD 60  
 QY 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVFPATDNLSSL 120  
 DB 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVFPATDNLSSL 120  
 QY 121 WGLKASILMKNQNDAAWEDLTRKETIDNNSVSPLOSQOORTWLIHWSLFPVFNHPKGR 180  
 DB 121 WGLKASILMKNQNDAAWEDLTRKETIDNNSVSPLOSQOORTWLIHWSLFPVFNHPKGR 180

QY 181 DNIIDFLYQPOYLNAIQMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTYK 240  
| | | | |  
Db 181 DNIIDFLYQPOYLNAIQMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTYK 240  
| | | | |  
QY 241 DPITEFVECLYVNFDFGAKKLRECESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300  
| | | | |  
Db 241 DPITEFVECLYVNFDFGAKKLRECESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300  
| | | | |  
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTK 360  
| | | | |  
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTK 360  
| | | | |  
QY 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396  
| | | | |  
Db 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396  
| | | | |  
RESULT 11  
ABB57306  
ID ABB57306 standard; protein; 445 AA.  
XX  
AC ABB57306;  
XX  
DT 07-MAR-2002 (first entry)  
XX  
DE Mouse ischaemic condition related protein sequence SEQ ID NO:856.  
XX  
XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;  
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.  
XX  
OS Mus musculus.  
XX  
PN WO20018188-A2.  
XX  
PD 22-NOV-2001.  
XX  
PF 18-MAY-2001; 2001WO-JP004192.  
XX  
PR 18-MAY-2000; 2000JP-00145977.  
XX  
PA (UYNI-) UNIV NTHON SCHOOL JURIDICAL PERSON.  
XX  
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;  
DR WPI; 2002-034733/04.  
DR N-PSDB; ABI99770.  
XX  
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
PT expression levels of particular genes defined in the specification or by  
PT determining the expression profile of a gene group comprising these  
PT genes.  
XX  
PS Claim 2; Page 2125-2127; 2690pp; English.  
XX  
XX The present invention describes a method for examining ischaemic  
CC conditions, comprising measuring the expression levels of particular  
CC genes (I) in a test sample or determining the expression profile of a  
CC gene group in the sample comprising genes selected from (I). The method  
CC is useful for examining the ischaemic condition (e.g. compressive  
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
CC expression levels of particular genes (ABI99202 to ABI99912, encoding the  
CC protein sequences in ABB57020 to ABB57374) or by determining the  
CC expression profile of a gene group comprising these genes. The expression  
CC levels or expression profiles produced by these genes are used as an  
CC indicator when screening for ischaemic condition-improving drugs or  
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR  
CC primers for a mouse ischaemic condition related sequence, which are used  
CC in the exemplification of the present invention  
XX  
SQ Sequence 445 AA;

Query Match 100.0%; Score 396; DB 5; Length 445;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MYDFAMDVYKLYSDDIPHALREKRTTVVAQLKQLAETETPIVKMFEDPETTRQMOSTRD 60  
| | | | |  
Db 50 MYDFAMDVYKLYSDDIPHALREKRTTVVAQLKQLAETETPIVKMFEDPETTRQMOSTRD 109  
| | | | |  
QY 61 GRMLFDYLDADKHGFRQYLDITLYRYAKFOYECGNYSAAEYLYFFRVLVPATDRLNALSSL 120  
| | | | |  
Db 110 GRMLFDYLDADKHGFRQYLDITLYRYAKFOYECGNYSAAEYLYFFRVLVPATDRLNALSSL 169  
| | | | |  
QY 121 WKGLASEIILMOWDAAMEDLTRKETIDNNSVSSPLOSLOORTWLHWSLFVFFNHPKGR 180  
| | | | |  
Db 170 WKGLASEIILMOWDAAMEDLTRKETIDNNSVSSPLOSLOORTWLHWSLFVFFNHPKGR 229  
| | | | |  
QY 181 DNIIDFLYQPOYLNAIQMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTYK 240  
| | | | |  
Db 230 DNIIDFLYQPOYLNAIQMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTYK 289  
| | | | |  
QY 241 DPITEFVECLYVNFDFGAKKLRECESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300  
| | | | |  
Db 290 DPITEFVECLYVNFDFGAKKLRECESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 349  
| | | | |  
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTK 360  
| | | | |  
Db 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTK 409  
| | | | |  
QY 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396  
| | | | |  
Db 410 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445  
| | | | |  
RESULT 12  
ABR39935  
ID ABR39935 standard; protein; 445 AA.  
XX  
AC ABR39935;  
XX  
DT 11-AUG-2003 (first entry)  
XX  
DE Human prostate selective polypeptide Pr327.  
XX  
KW Prostate; molecular marker; cancer; cytostatic; gene therapy; human.  
OS Homo sapiens.  
XX  
PN WO2003014298-A2.  
XX  
PD 20-FEB-2003.  
XX  
PF 02-AUG-2002; 2002WO-US024431.  
XX  
PR 03-AUG-2001; 2001US-0309470P.  
PR 30-OCT-2001; 2001US-0330747P.  
XX  
XX (ORIG-) ORIGENE TECHNOLOGIES INC.  
PA  
XX Sun Z, Li X, Jay G, Kovacs KF, Shu Y, Fan W;  
PI WPI; 2003-256562/25.  
XX  
DR N-PSDB; ACC47340.  
XX  
XX New polynucleotide, useful for preparing a composition for treating  
PT prostate disease, e.g., cancer.  
PT  
XX Claim 5; Page 150-152; 212pp; English.  
XX  
XX The invention relates to prostate selective polynucleotides and  
CC polypeptides. The polynucleotides are expressed in prostate and are  
CC useful as molecular markers, as drug targets, and for detecting,  
CC monitoring, preventing or treating diseases and conditions related to  
CC prostate, such as prostate cancers. The present sequence represents a  
CC prostate specific polypeptide

```
XX SQ Sequence 445 AA;
Query Match 100.0%; Score 396; DB 6; Length 445;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLOAETEPVVKMFPEDPETTRQMOSTRD 60
DB 50 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLOAETEPVVKMFPEDPETTRQMOSTRD 109
QY 61 GRMLFDYLDKKGFRQYELDTLYRYAKFOYECGNYSGAAYLYFFRVLVPATDNLSSL 120
DB 110 GRMLFDYLDKKGFRQYELDTLYRYAKFOYECGNYSGAAYLYFFRVLVPATDNLSSL 169
QY 121 WGLKASEILMNQNDAAWEDLTRKETIDNNSVSPLOSLOORTWLHWSLFFVFNHPKGR 180
DB 170 WGLKASEILMNQNDAAWEDLTRKETIDNNSVSPLOSLOORTWLHWSLFFVFNHPKGR 229
QY 181 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTK 240
DB 230 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTK 289
QY 241 DPITEFVECLYVNFDFDGAQKLRCECVSLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
DB 290 DPITEFVECLYVNFDFDGAQKLRCECVSLVNDFFLVACLEDFTENARLFIFETFCRIHQ 349
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
DB 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 409
QY 361 SLSFRSOMLAWNTEKKLNQNSRSEAPNWTQDSGFY 396
DB 410 SLSFRSOMLAWNTEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 13
ABU05160
ID ABU05160 standard; protein; 445 AA.
AC ABU05160;
DT 29-JAN-2003 (first entry)
DE Human expressed protein tag (EPT) #1826.
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX Homo sapiens.
XX OS
XX PN WO200278524-A2.
XX PD 10-OCT-2002.
XX PF 28-MAR-2002; 2002WO-US009671.
XX PR 28-MAR-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.
XX PR 01-OCT-2001; 2001US-0326370P.
XX PR 04-DEC-2001; 2001US-0336780P.
XX PR 20-FEB-2002; 2002US-0358985P.
XX PA (ZYCO-) ZYCOS INC.
XX PI Chicx RM, Tomlinson AJ, Urban RG;
XX DR WPI; 2003-040607/03.
XX
```

```
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX Example 2; SEQ ID NO 1826; 134pp; English.
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor, a
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide, is useful for treating cancer, or the antibody that binds to this
XX polypeptide, is useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: this sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 445 AA;

Query Match 100.0%; Score 396; DB 6; Length 445;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLOAETEPVVKMFPEDPETTRQMOSTRD 60
DB 50 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLOAETEPVVKMFPEDPETTRQMOSTRD 109
QY 61 GRMLFDYLDKKGFRQYELDTLYRYAKFOYECGNYSGAAYLYFFRVLVPATDNLSSL 120
DB 110 GRMLFDYLDKKGFRQYELDTLYRYAKFOYECGNYSGAAYLYFFRVLVPATDNLSSL 169
QY 121 WGLKASEILMNQNDAAWEDLTRKETIDNNSVSPLOSLOORTWLHWSLFFVFNHPKGR 180
DB 170 WGLKASEILMNQNDAAWEDLTRKETIDNNSVSPLOSLOORTWLHWSLFFVFNHPKGR 229
QY 181 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTK 240
DB 230 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTK 289
QY 241 DPITEFVECLYVNFDFDGAQKLRCECVSLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
DB 290 DPITEFVECLYVNFDFDGAQKLRCECVSLVNDFFLVACLEDFTENARLFIFETFCRIHQ 349
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
DB 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 409
QY 361 SLSFRSOMLAWNTEKKLNQNSRSEAPNWTQDSGFY 396
DB 410 SLSFRSOMLAWNTEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 14
ABU05163
ID ABU05163 standard; protein; 445 AA.
XX AC ABU05163;
XX DT 29-JAN-2003 (first entry)
XX DE Human expressed protein tag (EPT) #1829.
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX KW protease; protease inhibitor; transporter; cytoskeletal protein;
XX KW receptor; transcription factor; cancer; MHC;
```

KW	major histocompatibility complex; myeloma; colon cancer; gastric cancer; adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
KW	
XX	Homo sapiens.
XX	WO200278524-A2.
XX	
XX	10-OCT-2002.
XX	
XX	28-MAR-2002; 2002WO-US0009671.
XX	
XX	28-MAR-2001; 2001US-0279495P.
PR	21-MAY-2001; 2001US-0292544P.
PR	08-AUG-2001; 2001US-0310801P.
PR	01-OCT-2001; 2001US-0326370P.
PR	04-DEC-2001; 2001US-0336780P.
PR	20-FEB-2002; 2002US-0358985P.
XX	
XX	(ZYCO-) ZYCOS INC.
XX	
XX	Chicz RM, Tomlinson AJ, Urban RG;
XX	
XX	WPI; 2003-040607/03.
XX	
XX	New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.
XX	
XX	Example 2; SEQ ID NO 1829; 134pp; English.
XX	
XX	The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 445 AA;
XX	
XX	Query Match
XX	Best Local Similarity 100.0%; Score 396; DB 6; Length 445;
XX	Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MVDFAVDVKNLYSDDIPHALREKRTTVVAQLKQLAQETPIVKMFEDPETTRQMOSTRD 60
DB	50 MVDFAVDVKNLYSDDIPHALREKRTTVVAQLKQLAQETPIVKMFEDPETTRQMOSTRD 109
QY	61 GRMLFDYLDKGFQREYLDLTRYAKFQYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
DB	110 GRMLFDYLDKGFQREYLDLTRYAKFQYECGNYSGAAEYLYFFRVLVPATDRNALSSL 169
QY	121 WGLASEILMWNQDAWEDLTRKETIDNNSVSSPLOSQORTWLHWSLFFVFNHPKGR 180
DB	170 WGLASEILMWNQDAWEDLTRKETIDNNSVSSPLOSQORTWLHWSLFFVFNHPKGR 229
QY	181 DNIIDLFLYQPYLNAIQTCWCHILRYLTAVITNKDKVRKROVLKVLKQESYTYK 240
DB	230 DNIIDLFLYQPYLNAIQTCWCHILRYLTAVITNKDKVRKROVLKVLKQESYTYK 289
QY	241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300

Db	290 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 349
QY	301 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIEKTK 360
Db	350 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIEKTK 409
QY	361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db	410 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445
XX	
XX	RESULT 15
XX	ABU05157
ID	ABU05157 standard; protein; 445 AA.
XX	
XX	ABU05157;
XX	
XX	29-JAN-2003 (first entry)
XX	
XX	Human expressed protein tag (EPT) #1823.
XX	
XX	Translational profiling; expressed protein tag; EPT; kinase; phosphatase; protease; protease inhibitor; transporter; cytoskeletal protein; receptor; transcription factor; cancer; MHC;
KW	major histocompatibility complex; myeloma; colon cancer; gastric cancer; adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX	
XX	Homo sapiens.
XX	
XX	WO200278524-A2.
XX	
XX	10-OCT-2002.
XX	
XX	28-MAR-2002; 2002WO-US0009671.
XX	
XX	28-MAR-2001; 2001US-0279495P.
PR	21-MAY-2001; 2001US-0292544P.
PR	08-AUG-2001; 2001US-0310801P.
PR	01-OCT-2001; 2001US-0326370P.
PR	04-DEC-2001; 2001US-0336780P.
PR	20-FEB-2002; 2002US-0358985P.
XX	
XX	(ZYCO-) ZYCOS INC.
XX	
XX	Chicz RM, Tomlinson AJ, Urban RG;
XX	
XX	WPI; 2003-040607/03.
XX	
XX	New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.
XX	
XX	Example 2; SEQ ID NO 1823; 134pp; English.
XX	
XX	The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 445 AA;
XX	
XX	Query Match
XX	Best Local Similarity 100.0%; Score 396; DB 6; Length 445;
XX	Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MVDFAVDVKNLYSDDIPHALREKRTTVVAQLKQLAQETPIVKMFEDPETTRQMOSTRD 60
DB	50 MVDFAVDVKNLYSDDIPHALREKRTTVVAQLKQLAQETPIVKMFEDPETTRQMOSTRD 109
QY	61 GRMLFDYLDKGFQREYLDLTRYAKFQYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
DB	110 GRMLFDYLDKGFQREYLDLTRYAKFQYECGNYSGAAEYLYFFRVLVPATDRNALSSL 169
QY	121 WGLASEILMWNQDAWEDLTRKETIDNNSVSSPLOSQORTWLHWSLFFVFNHPKGR 180
DB	170 WGLASEILMWNQDAWEDLTRKETIDNNSVSSPLOSQORTWLHWSLFFVFNHPKGR 229
QY	181 DNIIDLFLYQPYLNAIQTCWCHILRYLTAVITNKDKVRKROVLKVLKQESYTYK 240
DB	230 DNIIDLFLYQPYLNAIQTCWCHILRYLTAVITNKDKVRKROVLKVLKQESYTYK 289
QY	241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300



SQ Sequence 445 AA;			Query Match			100.0%; Score 396; DB 6; Length 445;																																																					
Best Local Similarity 100.0%; Pred. No. 0;																																																											
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;																																																											
Qy	1	MVD	F	A	M	V	Y	K	N	L	S	D	D	I	P	H	A	L	R	E	K	T	T	V	V	A	Q	L	K	Q	A	E	T	E	P	I	V	K	M	F	E	D	P	E	T	T	R	Q	M	S	T	R	D	60					
Db	50	MVD	F	A	M	V	Y	K	N	L	S	D	D	I	P	H	A	L	R	E	K	T	T	V	V	A	Q	L	K	Q	A	E	T	E	P	I	V	K	M	F	E	D	P	E	T	T	R	Q	M	S	T	R	D	109					
Qy	61	GR	M	L	F	D	Y	L	A	D	K	H	G	F	R	O	E	Y	L	D	T	L	Y	R	A	K	F	O	Y	E	C	G	N	Y	S	G	A	A	E	Y	L	F	F	R	V	L	V	P	A	T	D	R	N	A	L	S	S	L	120
Db	110	GR	M	L	F	D	Y	L	A	D	K	H	G	F	R	O	E	Y	L	D	T	L	Y	R	A	K	F	O	Y	E	C	G	N	Y	S	G	A	A	E	Y	L	F	F	R	V	L	V	P	A	T	D	R	N	A	L	S	S	L	169
Qy	121	W	G	K	L	A	S	E	I	L	M	O	N	W	D	A	A	M	E	D	L	T	R	L	K	E	T	I	D	N	N	S	V	S	P	L	O	S	L	O	R	T	W	L	I	H	W	S	L	F	V	F	F	N	H	P	K	R	180
Db	170	W	G	K	L	A	S	E	I	L	M	O	N	W	D	A	A	M	E	D	L	T	R	L	K	E	T	I	D	N	N	S	V	S	P	L	O	S	L	O	R	T	W	L	I	H	W	S	L	F	V	F	F	N	H	P	K	R	229
Qy	181	D	N	I	I	D	L	F	L	Y	Q	P	Y	L	N	A	I	Q	T	M	C	P	H	I	L	R	V	L	T	T	A	V	I	T	N	K	D	V	R	K	R	Q	V	L	K	D	L	K	V	I	Q	E	S	Y	T	Y	K	240	
Db	230	D	N	I	I	D	L	F	L	Y	Q	P	Y	L	N	A	I	Q	T	M	C	P	H	I	L	R	V	L	T	T	A	V	I	T	N	K	D	V	R	K	R	Q	V	L	K	D	L	K	V	I	Q	E	S	Y	T	Y	K	289	
Qy	241	D	P	I	T	E	F	V	E	C	L	Y	N	F	D	F	D	G	A	Q	K	L	R	E	C	S	V	L	N	D	F	F	L	V	A	C	L	E	D	F	I	E	N	A	R	L	F	I	F	E	T	F	C	R	I	H	Q	300	
Db	290	D	P	I	T	E	F	V	E	C	L	Y	N	F	D	F	D	G	A	Q	K	L	R	E	C	S	V	L	N	D	F	F	L	V	A	C	L	E	D	F	I	E	N	A	R	L	F	I	F	E	T	F	C	R	I	H	Q	349	
Qy	301	C	I	S	I	N	M	L	A	D	K	N	M	T	P	E	E	A	E	R	I	V	N	L	I	R	N	A	R	L	D	A	K	I	D	S	K	L	G	H	V	M	G	N	N	A	V	S	P	Y	Q	O	V	I	E	K	T	K	360
Db	350	C	I	S	I	N	M	L	A	D	K	N	M	T	P	E	E	A	E	R	I	V	N	L	I	R	N	A	R	L	D	A	K	I	D	S	K	L	G	H	V	M	G	N	N	A	V	S	P	Y	Q	O	V	I	E	K	T	K	409
Qy	361	S	L	S	P	R	S	Q	M	L	A	M	N	I	E	K	L	N	O	N	S	R	S	E	A	P	N	W	A	T	O	D	S	G	F	Y	396																						
Db	410	S	L	S	P	R	S	Q	M	L	A	M	N	I	E	K	L	N	O	N	S	R	S	E	A	P	N	W	A	T	O	D	S	G	F	Y	445																						

Search completed: March 8, 2005, 20:22:44  
Job time : 168 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 8, 2005, 20:23:41 ; Search time 137 Seconds  
(without alignments)  
951.108 Million cell updates/sec

Title: US-10-783-415-4

Perfect score: 396

Sequence: 1 MVDPMADVKNLYSDDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table:

Gapco 60.0, Gapext 60.0

Searched: 1391452 seqs, 329044822 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	396	100.0	396	10	US-09-858-152A-2
2	396	100.0	396	10	US-09-858-152A-4
3	396	100.0	396	16	US-10-408-765A-2064
4	396	100.0	396	16	US-10-783-415-2
5	396	100.0	396	16	US-10-783-415-4
6	396	100.0	445	13	US-10-087-192-558
7	204	51.5	320	13	US-10-087-192-555
8	103	26.0	117	15	US-10-264-049-4238
9	45	11.4	45	9	US-09-864-761-42538
10	42	10.6	42	9	US-09-864-761-40729
11	42	10.6	42	9	US-09-864-761-42138
12	34	8.6	112	14	US-10-106-698-6920
13	12	3.0	437	15	US-10-424-599-266298

14	10	2.5	439	16	US-10-437-963-188307	Sequence 188307,
15	10	2.5	439	16	US-10-767-701-45301	Sequence 45301, A
16	10	2.5	444	16	US-10-437-963-188308	Sequence 188308,
17	9	2.3	9	14	US-10-022-066-57	Sequence 57, Appl
18	9	2.3	9	14	US-10-022-066-395	Sequence 395, App
19	9	2.3	130	15	US-10-424-599-217317	Sequence 217317,
20	9	2.3	147	15	US-10-425-114-49583	Sequence 49583, A
21	9	2.3	187	15	US-10-425-114-61814	Sequence 61814, A
22	8	2.0	401	15	US-10-424-599-243201	Sequence 243201, A
23	8	2.0	4092	15	US-10-369-493-22278	Sequence 22278, A
24	7	1.8	44	14	US-10-073-118-28	Sequence 28, Appl
25	7	1.8	45	14	US-10-289-135A-34	Sequence 34, Appl
26	7	1.8	57	15	US-10-424-599-192126	Sequence 192126,
27	7	1.8	70	14	US-10-029-386-30307	Sequence 30307, A
28	7	1.8	74	14	US-10-279-964-8	Sequence 8, Appli
29	7	1.8	75	15	US-10-424-599-240685	Sequence 240685,
30	7	1.8	81	15	US-10-291-172-276	Sequence 276, App
31	7	1.8	81	15	US-10-221-278-276	Sequence 276, App
32	7	1.8	81	15	US-10-468-125-12	Sequence 12, Appl
33	7	1.8	87	15	US-10-424-599-189674	Sequence 189674,
34	7	1.8	89	15	US-10-291-172-652	Sequence 652, App
35	7	1.8	89	15	US-10-221-278-652	Sequence 652, App
36	7	1.8	108	15	US-10-424-599-171619	Sequence 171619,
37	7	1.8	122	16	US-10-437-963-191022	Sequence 191022,
38	7	1.8	131	15	US-10-424-599-190181	Sequence 190181,
39	7	1.8	133	9	US-09-870-844-6	Sequence 6, Appli
40	7	1.8	146	16	US-10-437-963-192808	Sequence 192808,
41	7	1.8	179	14	US-10-316-253-295	Sequence 295, App
42	7	1.8	181	10	US-09-847-208-154	Sequence 154, App
43	7	1.8	181	14	US-10-316-253-297	Sequence 297, App
44	7	1.8	181	15	US-10-240-240A-10	Sequence 10, Appl
45	7	1.8	182	9	US-09-374-671-89	Sequence 89, Appl

## ALIGNMENTS

## RESULT 1

US-09-858-152A-2  
; Sequence 2, Application US/09858152A  
; Publication No. US2003004419A1  
; GENERAL INFORMATION:  
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE  
; SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES  
; APPLICANT: Marchetti, Antonio  
; APPLICANT: Buttitta, Fiama  
; APPLICANT: Smith, Gilbert H.  
; APPLICANT: Callahan, Robert  
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6  
; FILE REFERENCE: 4239-59122  
; CURRENT APPLICATION NUMBER: US/09/858,152A  
; CURRENT FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Murine INT6  
US-09-858-152A-2

Query Match	100.0%	Score 396;	DB 10;	Length 396;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 396;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MVDPMADVKNLYSDDIPHA	REKRTTVAQKQLOAETEP	IVKMFEDPPTTQMQSTRD 60
Db	1	MVDPMADVKNLYSDDIPHA	REKRTTVAQKQLOAETEP	IVKMFEDPPTTQMQSTRD 60
Qy	61	GRMLFDYLDKGGFRQBYDL	TRYAKFQYECGNYGAAEY	LYFRVLVLPATDNNALSSL 120
Db	61	GRMLFDYLDKGGFRQBYDL	TRYAKFQYECGNYGAAEY	LYFRVLVLPATDNNALSSL 120
Qy	121	WGKLASEILMQNDAMEDL	TRLKETDNNSSVPSLQ	ORTWLIIHWSLFVFNHPKGR 180

```
Db 121 WGKLASEILMNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIIHWSLFFVFNHPKGR 180
Qy 181 DNIIDFLYQOYLNAIQTCWPHILRYLTAVITNKDVKRRQVLKDLVKVIOQESYTK 240
Db 181 DNIIDFLYQOYLNAIQTCWPHILRYLTAVITNKDVKRRQVLKDLVKVIOQESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Qy 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 2
US-09-858-152A-4
; Sequence 4, Application US/09858152A
; Publication No. US2003004419A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE
; APPLICANT: SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
; CURRENT APPLICATION NUMBER: US/09/858,152A
; CURRENT FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-152A-4

Query Match 100.0%; Score 396; DB 10; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMOSTRD 60
Db 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMOSTRD 60
Qy 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVDPATRNALSSL 120
Db 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVDPATRNALSSL 120
Qy 121 WGKLASEILMNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIIHWSLFFVFNHPKGR 180
Db 121 WGKLASEILMNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIIHWSLFFVFNHPKGR 180
Qy 181 DNIIDFLYQOYLNAIQTCWPHILRYLTAVITNKDVKRRQVLKDLVKVIOQESYTK 240
Db 181 DNIIDFLYQOYLNAIQTCWPHILRYLTAVITNKDVKRRQVLKDLVKVIOQESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Qy 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 3
US-10-408-765A-2064
; Sequence 2064, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale B.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2064
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2064

Query Match 100.0%; Score 396; DB 16; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMOSTRD 60
Db 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMOSTRD 60
Qy 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVDPATRNALSSL 120
Db 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVDPATRNALSSL 120
Qy 121 WGKLASEILMNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIIHWSLFFVFNHPKGR 180
Db 121 WGKLASEILMNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIIHWSLFFVFNHPKGR 180
Qy 181 DNIIDFLYQOYLNAIQTCWPHILRYLTAVITNKDVKRRQVLKDLVKVIOQESYTK 240
Db 181 DNIIDFLYQOYLNAIQTCWPHILRYLTAVITNKDVKRRQVLKDLVKVIOQESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Qy 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 4
US-10-783-415-2
; Sequence 2, Application US/10783415
; Publication No. US20040141918A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
```

```
Db 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 3
US-10-408-765A-2064
; Sequence 2064, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale B.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2064
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2064

Query Match 100.0%; Score 396; DB 16; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMOSTRD 60
Db 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMOSTRD 60
Qy 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVDPATRNALSSL 120
Db 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVDPATRNALSSL 120
Qy 121 WGKLASEILMNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIIHWSLFFVFNHPKGR 180
Db 121 WGKLASEILMNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIIHWSLFFVFNHPKGR 180
Qy 181 DNIIDFLYQOYLNAIQTCWPHILRYLTAVITNKDVKRRQVLKDLVKVIOQESYTK 240
Db 181 DNIIDFLYQOYLNAIQTCWPHILRYLTAVITNKDVKRRQVLKDLVKVIOQESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Qy 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSOMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 4
US-10-783-415-2
; Sequence 2, Application US/10783415
; Publication No. US20040141918A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
```

; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6  
; FILE REFERENCE: 4239-59122  
; CURRENT APPLICATION NUMBER: US/10/783,415  
; CURRENT FILING DATE: 2004-02-19  
; PRIOR APPLICATION NUMBER: 09/858,152  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Murine INT6  
US-10-783-415-2

Query Match 100.0%; Score 396; DB 16; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 60  
DB 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 60  
  
QY 61 GRMLFDYLDADKHGFRQEVLDLTRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120  
DB 61 GRMLFDYLDADKHGFRQEVLDLTRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120  
  
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLIIHWSLFFVFNHPKGR 180  
DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLIIHWSLFFVFNHPKGR 180  
  
QY 181 DNIIDFLYQPOLNAIQTMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240  
DB 181 DNIIDFLYQPOLNAIQTMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240  
  
QY 241 DPITEFVECLYVNFDFDGAQKLRCEESLVNDFLVACLEDFIENARLFIFETFCRIHQ 300  
DB 241 DPITEFVECLYVNFDFDGAQKLRCEESLVNDFLVACLEDFIENARLFIFETFCRIHQ 300  
  
QY 301 CISINMLADKLNMTPEEAERWIVNIRNARDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
DB 301 CISINMLADKLNMTPEEAERWIVNIRNARDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
  
QY 361 SLFSRQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396  
DB 361 SLFSRQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

## RESULT 5

US-10-783-415-4  
; Sequence 4, Application US/10783415  
; Publication No. US20040141918A1  
; GENERAL INFORMATION:  
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS  
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND  
; APPLICANT: HUMAN SERVICES  
; APPLICANT: Marchetti, Antonio  
; APPLICANT: Buttitta, Flamma  
; APPLICANT: Smith, Gilbert H.  
; APPLICANT: Callahan, Robert  
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6  
; FILE REFERENCE: 4239-59122  
; CURRENT APPLICATION NUMBER: US/10/783,415  
; CURRENT FILING DATE: 2004-02-19  
; PRIOR APPLICATION NUMBER: 09/858,152  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-783-415-4

Query Match 100.0%; Score 396; DB 16; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 60  
DB 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 60  
  
QY 61 GRMLFDYLDADKHGFRQEVLDLTRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120  
DB 61 GRMLFDYLDADKHGFRQEVLDLTRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120  
  
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLIIHWSLFFVFNHPKGR 180  
DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLIIHWSLFFVFNHPKGR 180  
  
QY 181 DNIIDFLYQPOLNAIQTMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240  
DB 181 DNIIDFLYQPOLNAIQTMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240  
  
QY 241 DPITEFVECLYVNFDFDGAQKLRCEESLVNDFLVACLEDFIENARLFIFETFCRIHQ 300  
DB 241 DPITEFVECLYVNFDFDGAQKLRCEESLVNDFLVACLEDFIENARLFIFETFCRIHQ 300  
  
QY 301 CISINMLADKLNMTPEEAERWIVNIRNARDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
DB 301 CISINMLADKLNMTPEEAERWIVNIRNARDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
  
QY 361 SLFSRQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396  
DB 361 SLFSRQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

## RESULT 6

US-10-087-192-558  
; Sequence 558, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: CANCER  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 558  
; LENGTH: 445  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-087-192-558

Query Match 100.0%; Score 396; DB 13; Length 445;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 60  
DB 50 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 109  
  
QY 61 GRMLFDYLDADKHGFRQEVLDLTRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120  
DB 110 GRMLFDYLDADKHGFRQEVLDLTRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 169  
  
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLIIHWSLFFVFNHPKGR 180  
DB 170 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLIIHWSLFFVFNHPKGR 229

Qy 181 DNIIDFLYQYQYLNAIQMCPHILRYLTAVITNKQVRKROVLKOLVKVIOQESYTYK 240  
Db 230 DNIIDFLYQYQYLNAIQMCPHILRYLTAVITNKQVRKROVLKOLVKVIOQESYTYK 289  
Qy 241 DPITEFVECLVNFDFGAQKLRCECSVLVNDFFLVACLEDFFIENARLFIETFCRIHQ 300  
Db 290 DPITEFVECLVNFDFGAQKLRCECSVLVNDFFLVACLEDFFIENARLFIETFCRIHQ 349  
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360  
Db 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 409  
Qy 361 SLSFRSOMLANNIEKLNQNSRSEAPNWTQDSGFY 396  
Db 410 SLSFRSOMLANNIEKLNQNSRSEAPNWTQDSGFY 445

RESULT 7  
US-10-087-192-555  
; Sequence 555, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE OF INVENTION: CANCER  
; FILE REFERENCE: 52945200122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 555  
; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-087-192-555

Query Match 51.5%; Score 204; DB 13; Length 320;  
Best Local Similarity 100.0%; Pred. No. 2.8e-186;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 193 YLNAIQMCPHILRYLTAVITNKQVRKROVLKOLVKVIOQESYTYKDPITEFVECLYV 252  
Db 117 YLNAIQMCPHILRYLTAVITNKQVRKROVLKOLVKVIOQESYTYKDPITEFVECLYV 176  
Qy 253 NFDFDGAQKLRCECSVLVNDFFLVACLEDFFIENARLFIETFCRIHQCSINNMLADKLN 312  
Db 177 NFDFDGAQKLRCECSVLVNDFFLVACLEDFFIENARLFIETFCRIHQCSINNMLADKLN 236  
Qy 313 MTPPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETKXLSFRSOMLANN 372  
Db 237 MTPPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETKXLSFRSOMLANN 296  
Qy 373 IEKLNQNSRSEAPNWTQDSGFY 396  
Db 297 IEKLNQNSRSEAPNWTQDSGFY 320

RESULT 8  
US-10-264-049-4238  
; Sequence 4238, Application US/10264049  
; Publication No. US20040000579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P413P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569

; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 4238  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (4)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-049-4238  
Query Match 26.0%; Score 103; DB 15; Length 117;  
Best Local Similarity 100.0%; Pred. No. 4e-90;  
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 126 SEILMQNWDAAEDLTRELKETIDNNSSVSSPQSLQORTWLHWSLFFVFNHPKGRDNIID 185  
Db 5 SEILMQNWDAAEDLTRELKETIDNNSSVSSPQSLQORTWLHWSLFFVFNHPKGRDNIID 64  
Qy 186 LFLYQPOYLNAIQMCPHILRYLTAVITNKQVRKROVLKDL 228  
Db 65 LFLYQPOYLNAIQMCPHILRYLTAVITNKQVRKROVLKDL 107  
RESULT 9  
US-09-864-761-42538  
; Sequence 42538, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 42538  
; LENGTH: 45  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AP001331.1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.3  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.8  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 12  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.7  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 24  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 11  
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 10  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.8  
; OTHER INFORMATION: EST\_HUMAN HIT: A1718630.1, EVALUE 3.00e-18  
; OTHER INFORMATION: SWISSPROT HIT: Q64252, EVALUE 2.00e-19  
US-09-864-761-42538

Query Match 11.4%; Score 45; DB 9; Length 45;  
Best Local Similarity 100.0%; Pred. No. 4.7e-35; Indels 0; Gaps 0;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 340 GHVVGNNVSPYQOVIETKSLSPRSQMLAMNIEKLNQNSRSE 384  
Db 1 GHVVGNNVSPYQOVIETKSLSPRSQMLAMNIEKLNQNSRSE 45

RESULT 10  
US-09-864-761-40729  
; Sequence 40729, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 40729  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC020991.3  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.93  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8  
; OTHER INFORMATION: EST\_HUMAN HIT: BF700473.1, EVALUE 7.00e-17  
; OTHER INFORMATION: SWISSPROT HIT: Q64252, EVALUE 6.00e-18  
US-09-864-761-40729

Query Match 10.6%; Score 42; DB 9; Length 42;  
Best Local Similarity 100.0%; Pred. No. 3.3e-32; Indels 0; Gaps 0;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 193 YLNAIQTMCPHILRYLTAVITNKDVKRRQVLKDLVKVIOQ 234  
Db 1 YLNAIQTMCPHILRYLTAVITNKDVKRRQVLKDLVKVIOQ 42

RESULT 11  
US-09-864-761-42138  
; Sequence 42138, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42138
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP001331.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
; OTHER INFORMATION: EST HUMAN HIT: BF700473.1, EVALUAE 7.00e-17
; OTHER INFORMATION: SWISSPROT HIT: Q64252, EVALUAE 6.00e-18
US-09-864-761-42138

Query Match 10.6%; Score 42; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 3.3e-32;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 YLNAIQTMCPHILRYLTAVITNKDVKRQVLKDLVKVIQQ 234
Db 1 YLNAIQTMCPHILRYLTAVITNKDVKRQVLKDLVKVIQQ 42

RESULT 12
US-10-106-698-6920
; Sequence 6920, Application US/101066598
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6920
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)\_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (4)\_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (13)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (29)\_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6920

Query Match 8.6%; Score 34; DB 14; Length 112;
Best Local Similarity 100.0%; Pred. No. 3.4e-24;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDPMDDVYKNLYSDDIPHALREKRTTVVAQLKQ 34
Db 70 MVDPMDDVYKNLYSDDIPHALREKRTTVVAQLKQ 103

RESULT 13
US-10-424-599-266298
; Sequence 266298, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 266298
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_82489C.1.pep
US-10-424-599-266298

Query Match 3.0%; Score 12; DB 15; Length 437;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 316 EEAERWIVNLIR 327
Db 370 EEAERWIVNLIR 381

RESULT 14
US-10-437-963-188307
; Sequence 188307, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963



; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 188307  
; LENGTH: 439  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_84924C.1.pep  
US-10-437-963-188307

Query Match 2.5%; Score 10; DB 16; Length 439;  
Best Local Similarity 100.0%; Pred. No. 0.96;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 285 ENARLFIFET 294  
Db 339 ENARLFIFET 348

RESULT 15  
US-10-767-701-45301  
; Sequence 45301, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 45301  
; LENGTH: 439  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C1020\_1.pep  
US-10-767-701-45301

Query Match 2.5%; Score 10; DB 16; Length 439;  
Best Local Similarity 100.0%; Pred. No. 0.96;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 285 ENARLFIFET 294  
Db 339 ENARLFIFET 348

Search completed: March 8, 2005, 20:38:23  
Job time : 138 secs

**THIS PAGE BLANK (USSTO)**

## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 19:47:46 ; Search time 172 Seconds

(without alignments)  
1178.973 Million cell updates/sec

**Title:** US-10-783-415-4

Perfect score: 2071  
Sequence: 1 MVDFAMDVYKNLYSDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 03:\*

```
1: uniprot_sprot:*
2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query	Score	Query			DB	ID	Description
			Match	Length				
1	2071	100.0	396	2	Q9BRV2		Q9brv2	homo sapien
2	2071	100.0	445	1	IF36_HUMAN		P60228	homo sapien
3	2071	100.0	445	1	IF36_MOUSE		P60229	mus musculus
4	2071	100.0	445	2	Q641X8		Q641x8	rattus norv
5	2067	99.8	445	2	Q8WVK4		Q8wvk4	homo sapien
6	2064	99.7	445	2	Q61AX5		Q61ax5	homo sapien
7	2006.5	96.9	446	2	Q6P7L9		Q6p7l9	xenopus tro
8	1992.5	96.2	423	2	Q7ZXA5		Q7zxa5	xenopus lae
9	1981.5	95.7	446	2	Q918W7		Q918w7	xenopus lae
10	1952.5	94.3	446	2	Q6P7X8		Q6p7x8	brachydanio
11	1947.5	94.0	446	2	Q6DR11		Q6dr11	brachydanio
12	1946	94.0	422	2	Q9CT23		Q9ct23	mus musculus
13	1783.5	86.1	448	2	Q6TH16		Q6th16	brachydanio
14	1220.5	58.9	435	1	IF36_DROME		Q77410	drosophila g
15	1165.5	56.3	462	2	Q7Q1L8		Q7q1l8	anopheles g
16	1081.5	52.2	439	2	Q8GV10		Q8gv10	oryza sativ
17	1066	51.5	372	2	Q8MR88		Q8mr88	drosophila
18	1049.5	50.7	441	2	Q9C5Z3		Q9c5z3	arabidopsis
19	1048.5	50.6	441	2	Q9M4T7		Q9m4t7	arabidopsis
20	1011.5	48.0	415	2	Q6SW36		Q6sw36	oryza sativ
21	994.5	48.0	432	1	IF36_CAEEL		Q61820	caenorhabdi
22	990.5	47.8	397	2	Q6SW37		Q6sw37	oryza sativ
23	927.5	44.8	436	2	Q6Z475		Q6z475	oryza sativ
24	903	43.6	418	2	Q9M2L8		Q9m2l8	arabidopsis
25	791.5	38.2	443	2	Q78519		Q78519	neurospora
26	781	37.7	421	2	Q6CQ02		Q6cq02	varrowia li
27	775	37.4	501	1	IF36_SCHPO		Q94513	schizosacch
28	553	26.7	119	2	Q8BN66		Q8bn66	mus musculus
29	424	20.5	527	2	Q7RQ22		Q7rq22	plasmodium
30	413	19.9	517	2	Q81315		Q81315	leishmania
31	396.5	19.1	405	2	Q9NNE7		Q9nee7	leishmania

Qy 1 MVD FAMD VYKNLYSDDI PHALREKRTTVVAQLKQLQAETETPIVKMFEDPETTRQMQSTRD 60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

\_\_\_\_\_

Db 1 MVDAMDVYKNLYSDIPHALREKRTTVVQAQLKQIQAETPIVKMFDEPPTTQWQSTRD 60  
 Qy 61 GRMLFDYLDADKHGRFQEQEYLDLYRYAKFQYCGNYSGAAYLYFPRVLVPATDRLNLSL 120  
 Db 61 GRMLFDYLDADKHGRFQEQEYLDLYRYAKFQYCGNYSGAAYLYFPRVLVPATDRLNLSL 120  
 Qy 121 WGLKASEILMQWDAAMEDTLRLKETIDNNSSVSSPLOSLOQRTWLHWSLFFVFNHPKGR 180  
 Db 121 WGLKASEILMQWDAAMEDTLRLKETIDNNSSVSSPLOSLOQRTWLHWSLFFVFNHPKGR 180  
 Qy 181 DNIIDLFLYQYQYLNALQTMCPHILRYLTAVITNKQVRRQVLKDIKVKVIQESYTYK 240  
 Db 181 DNIIDLFLYQYQYLNALQTMCPHILRYLTAVITNKQVRRQVLKDIKVKVIQESYTYK 240  
 Qy 241 DPTEFVECLYVNFDFDGAQKLCESVNVNDFLVAACLEDFTENARLFIFETPCRHHQ 300  
 Db 241 DPTEFVECLYVNFDFDGAQKLCESVNVNDFLVAACLEDFTENARLFIFETPCRHHQ 300  
 Qy 301 CISINMLADKLNMTPEEARMTVNLIRNARDLDAKIDSKLGHVVMGNNVSPVQOVIETK 360  
 Db 301 CISINMLADKLNMTPEEARMTVNLIRNARDLDAKIDSKLGHVVMGNNVSPVQOVIETK 360  
 Qy 361 SLSPRSQMLAMNIEKLNQNSRSEAPNATQDSGFY 396  
 Db 361 SLSPRSQMLAMNIEKLNQNSRSEAPNATQDSGFY 396

RESULT 2  
 IF36 HUMAN STANDARD; PRT: 445 AA.  
 AC P60228; O43902; Q64058; Q64059; Q64252;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Eukaryotic translation initiation factor 3 subunit 6 (eIF-3 p48)  
 DE (eIF3e) (Viral integration site protein INT-6 homolog).  
 GN Name=eIF3e; Synonym=INT6;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 280-289 AND 427-436.  
 RC TISSUE=Liver;  
 RX MEDLINE=97442403; PubMed=9295280; DOI=10.1074/jbc.272.38.23477;  
 RA Asano K., Merrick W.C., Hershey J.W.B.;  
 RT "The translation initiation factor eIF3-p48 subunit is encoded by int-6, a site of frequent integration by the mouse mammary tumor virus genome.";  
 RL J. Biol. Chem. 272:23477-23480(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=98066777; PubMed=9403073; DOI=10.1006/geno.1997.4996;  
 RA Miyazaki S., Inatani A., Ballard L., Marchetti A., Buttitta F., Albertsen H., Nevanlinna H.A., Gallahan D., Callahan R.;  
 RT "The chromosome location of the human homolog of the mouse mammary tumor-associated gene INT6 and its status in human breast carcinomas";  
 RL Genomics 46:155-158(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96337990; PubMed=8688078;  
 RA Desbois C., Rousset R., Bantignies F., Jalinot P.;  
 RT "Exclusion of Int-6 from PML nuclear bodies by binding to the HTLV-I Tax oncoprotein.";  
 RL Science 273:951-953(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=12386384;  
 RA Neuvart C., Jin D.-Y., Semmes O.J., Diella F., Callahan R., Jeang K.-T.;  
 RT "Divergent subcellular locations of HTLV-I Tax and Int-6: a contrast between in vitro protein-protein binding and intracellular protein colocalization.";  
 RL J. Biomed. Sci. 4:229-234(1997).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow, Brain, and Muscle;  
 RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T.I., Max S.I., Wang J., Hsieh P., Stapleton L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP INTERACTION WITH TRIM27.  
 RX PubMed=10504338;  
 RA Morris-Desbois C., Bochar V., Reynaud C., Jalinot P.;  
 RT "Interaction between the Ret finger protein and the Int-6 gene product and co-localisation into nuclear bodies.";  
 RL Cell Sci. 112:3331-3342(1999).  
 CC -!- FUNCTION: Binds to the 40S ribosome and promotes the binding of methionyl-tRNAi and mRNA.  
 CC -!- SUBUNIT: eIF-3 is composed of at least 12 different subunits.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.  
 CC -!- SIMILARITY: Belongs to the eIF3e family.  
 CC -!- SIMILARITY: Contains 1 PCI domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U54562; AAC51760.1; --  
 CC EMBL; U94174; AAC51917.1; --  
 CC EMBL; U94162; AAC51917.1; JOINED.  
 CC EMBL; U94163; AAC51917.1; JOINED.  
 CC EMBL; U94164; AAC51917.1; JOINED.  
 CC EMBL; U94165; AAC51917.1; JOINED.  
 CC EMBL; U94166; AAC51917.1; JOINED.  
 CC EMBL; U94167; AAC51917.1; JOINED.  
 CC EMBL; U94168; AAC51917.1; JOINED.  
 CC EMBL; U94169; AAC51917.1; JOINED.  
 CC EMBL; U94170; AAC51917.1; JOINED.  
 CC EMBL; U94171; AAC51917.1; JOINED.  
 CC EMBL; U94172; AAC51917.1; JOINED.  
 CC EMBL; U94173; AAC51917.1; JOINED.  
 CC EMBL; U94175; AAC51919.1; --  
 CC EMBL; U62962; AAB58251.1; --  
 CC EMBL; U85947; AAB88873.1; --  
 CC EMBL; BC000734; AAH00734.1; --  
 CC EMBL; BC008419; AAH08419.1; --  
 CC EMBL; BC016706; AAH16706.1; --  
 CC EMBL; BC021679; AAH21679.1; --  
 CC IntAct; P60228; --

```
DR Genew; HGNC:3277; EIF3S6.
DR H-InvDB; HIX0007722; -.
DR Reactome; P60228; -.
DR MIT; 602210; -.
DR InterPro; IPR000717; PCI.
DR Direct protein sequencing; Initiation factor; Protein biosynthesis.
DR KQ SEQUENCE 445 AA; 52220 MW; A5368651DDDDDDC CRC64;

Query Match 100.0%; Score 2071; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 5.8e-145;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVAQLKQLQAETPIVPMFEDPETTRQMOSTRD 60
DB 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVAQLKQLQAETPIVPMFEDPETTRQMOSTRD 109
QY 61 GRMLFDYLDKXGFRQYEDLTLYRYAKFQYECGNYSGAAEYLYFFRVLVPAIDRNALSSL 120
DB 110 GRMLFDYLDKXGFRQYEDLTLYRYAKFQYECGNYSGAAEYLYFFRVLVPAIDRNALSSL 169
QY 121 WGLKASEILMOWDAAMEDLTRKETIDNNSVSSPQSQOQRTWLIHWSLFVFNHPKGR 180
DB 170 WGLKASEILMOWDAAMEDLTRKETIDNNSVSSPQSQOQRTWLIHWSLFVFNHPKGR 229
QY 181 DNIIDFLYQPOYLNAIQTCMCHILRYLTAVITNKDVKRQVLKDLVKVYQESYTYK 240
DB 230 DNIIDFLYQPOYLNAIQTCMCHILRYLTAVITNKDVKRQVLKDLVKVYQESYTYK 289
QY 241 DPTFVECLYNFDPGAKKLRECVSLVNDFFLVACLEDPIENARLPIFETFCRIHQ 300
DB 290 DPTFVECLYNFDPGAKKLRECVSLVNDFFLVACLEDPIENARLPIFETFCRIHQ 349
QY 301 CTSINMLADKLNMTPEASRWIINLRNARLDKIDSKLGHVVMGNVSPQVQVIEKTK 360
DB 350 CTSINMLADKLNMTPEASRWIINLRNARLDKIDSKLGHVVMGNVSPQVQVIEKTK 409
QY 361 SLSFRSOMLAMIETKLNQNSRSEAPNWTQDSGFY 396
DB 410 SLSFRSOMLAMIETKLNQNSRSEAPNWTQDSGFY 445

RESULT 3
ID IF36 MOUSE STANDARD; PRT: 445 AA.
AC P60229; Q64058; Q64059; Q64252;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Eukaryotic translation initiation factor 3 subunit 6 (eIF-3 p48)
DE (eIF3e) (Mammary tumor-associated protein INT-6) (Viral integration
DE site protein INT-6) (MMTV integration site 6);
GN Name=EIF36; Synonyms=Int6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN [2]
RN [3]
SEQUENCE FROM N.A.
RX MEDLINE=95156630; PubMed=7853537;
RA Marchetti A., Buttitta F., Miyazaki S., Gallahan D., Smith G.H.,
RA Callahan R.;
RT "Int-6, a highly conserved, widely expressed gene, is mutated by mouse
RT mammary tumor virus in mammary preneoplasia.";
RL J. Virol. 69:1932-1938 (1995).
RN [2]
RN [3]
SEQUENCE FROM N.A.
RX MEDLINE=97405883; PubMed=9260927;
RA Diella F., Levi G., Callahan R.;
RT "Characterization of the INT6 mammary tumor gene product.";
RT DNA Cell Biol. 16:839-847 (1997).
RN [3]
RN [4]
RN [5]
REVIEWS TO N-TERMINUS.
RA Callahan R.;
```

```
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II; TISSUE=Mammary gland;
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins S.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Musny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalish D.E.,
RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Binds to the 40S ribosome and promotes the binding of
CC methionyl-tRNAi and mRNA.
CC -!- SUBUNIT: eIF-3 is composed of at least 12 different subunits.
CC Interacts with TRIM27 (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
CC -!- DISEASE: Int-6 serves as a site for viral integration of mouse
CC mammary tumor virus (MMTV) in mammary tumors.
CC -!- SIMILARITY: Belongs to the EIF36 family.
CC -!- SIMILARITY: Contains 1 PCI domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S75221; AAC00046.1; -.
CC EMBL; S75223; AAC00047.1; -.
CC EMBL; BC029177; AAH29177.1; -.
CC MGD; MGI:99257; Eif3e6.
CC InterPro; IPR000717; PCI.
KW Initiation factor; Protein biosynthesis; Proto-oncogene.
FT VARIANT 158 158 180
FT VARIANT 158 180
FT VARIANT 159 445
FT VARIANT 318 365
FT VARIANT 366 445
FT VARIANT 445 AA; 52220 MW; A5368651DDDDDDC CRC64;
SQ SEQUENCE 445 AA; 52220 MW; A5368651DDDDDDC CRC64;

Query Match 100.0%; Score 2071; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 5.8e-145;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVAQLKQLQAETPIVPMFEDPETTRQMOSTRD 60
DB 50 MVDFAVDVYKNLYSDDIIPHALREKRTTVAQLKQLQAETPIVPMFEDPETTRQMOSTRD 109
QY 61 GRMLFDYLDKXGFRQYEDLTLYRYAKFQYECGNYSGAAEYLYFFRVLVPAIDRNALSSL 120
DB 110 GRMLFDYLDKXGFRQYEDLTLYRYAKFQYECGNYSGAAEYLYFFRVLVPAIDRNALSSL 169
```

```

QY 121 WGLASEILMNQWDAAMEDTLRLKETIDNNSSVSPLOSLQOQRTWLIIHWSLFVFFNHPKGR 180
DB 170 WGLASEILMNQWDAAMEDTLRLKETIDNNSSVSPLOSLQOQRTWLIIHWSLFVFFNHPKGR 229
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQBSYTYK 240
DB 230 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQBSYTYK 289
QY 241 DPITFEVCLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 300
DB 290 DPITFEVCLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 349
QY 301 CISINMLADKLNMTPEEAERWIVNLRNARLDKIDSGLGHVVMGNNAVSPYQOVIETK 360
DB 350 CISINMLADKLNMTPEEAERWIVNLRNARLDKIDSGLGHVVMGNNAVSPYQOVIETK 409
QY 361 SLSFRSOMLANNIEKLNQNSRSEAPNATODSGFY 396
DB 410 SLSFRSOMLANNIEKLNQNSRSEAPNATODSGFY 445

RESULT 4
Q641X8
ID Q641X8 PRELIMINARY; PRT; 445 AA.
AC Q641X8;
DT 25-OCT-2004 (TremBLrel. 28, Created)
DT 25-OCT-2004 (TremBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TremBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Testis;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC Tissue=Testis;
RA Director MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC082087; AAH82087.1; -
KW Hypothetical protein.
SQ SEQUENCE 445 AA; 52220 MW; A5368651DDDDDDOC CRC64;

Query Match 100.0%; Score 2071; DB 2; Length 445;
Best Local Similarity 100.0%; Pred. No. 5.8e-145;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDPMVDVYKNLYSDDIPHALREKRTTVVAQLKQLAQETPIVKMFPEDPTTRQMSTRD 60
DB 50 MVDPMVDVYKNLYSDDIPHALREKRTTVVAQLKQLAQETPIVKMFPEDPTTRQMSTRD 109

```

```

QY 61 GRMLFDYLADKHGFRQEYLDLTRYAKFOYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
DB 110 GRMLFDYLADKHGFRQEYLDLTRYAKFOYECGNYSGAAEYLYFFRVLPATDRNALSSL 169
QY 121 WGLASEILMNQWDAAMEDTLRLKETIDNNSSVSPLOSLQOQRTWLIIHWSLFVFFNHPKGR 180
DB 170 WGLASEILMNQWDAAMEDTLRLKETIDNNSSVSPLOSLQOQRTWLIIHWSLFVFFNHPKGR 229
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQBSYTYK 240
DB 230 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQBSYTYK 289
QY 241 DPITFEVCLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 300
DB 290 DPITFEVCLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 349
QY 301 CISINMLADKLNMTPEEAERWIVNLRNARLDKIDSGLGHVVMGNNAVSPYQOVIETK 360
DB 350 CISINMLADKLNMTPEEAERWIVNLRNARLDKIDSGLGHVVMGNNAVSPYQOVIETK 409
QY 361 SLSFRSOMLANNIEKLNQNSRSEAPNATODSGFY 396
DB 410 SLSFRSOMLANNIEKLNQNSRSEAPNATODSGFY 445

```

## RESULT 5

```

Q8WVK4
ID Q8WVK4 PRELIMINARY; PRT; 445 AA.
AC Q8WVK4;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Mammary tumor integration site 6 (Oncogene homolog).
GN Name=E1F3S6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC Tissue=Lung;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017887; AAH17887.1; -
DR PFAM; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
SQ SEQUENCE 445 AA; 52248 MW; AF0AD651DAAA8D0C CRC64;

```



```
Query Match          99.8%; Score 2067; DB 2; Length 445;
Best Local Similarity 99.7%; Pred. No. 1.1e-144;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDI PHALREKRTTVVAQLKQLQAEETPIVKMFEDPETTRQMOSTRD 60
DB 50 MVDFAVDVYKNLYSDDI PHALREKRTTVVAQLKQLQAEETPIVKMFEDPETTRQMOSTRD 109
QY 61 GRMLFDYLDKKGFRQYEDTLRYAKFOYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
DB 110 GRMLFDYLDKKGFRQYEDTLRYAKFOYECGNYSGAAEYLYFFRVLPATDRNALSSL 169
QY 121 WGLKASEILMQNWDAAEMDLTKETIDNNSSVSPLOSLQORTWLHWSLFFVFNHPKGR 180
DB 170 WGLKASEILMQNWDAAEMDLTKETIDNNSSVSPLOSLQORTWLHWSLFFVFNHPKGR 229
QY 181 DNIIDLFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRROVLKDLVKVIOQESYTYK 240
DB 230 DNIIDLFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRROVLKDLVKVIOQESYTYK 289
QY 241 DPITEFVECLYVNFDFDGAQKLCRECEVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB 290 DPITEFVECLYVNFDFDGAQKLCRECEVLVNDFFLVACLEDFIENARLFIETFCRIHQ 349
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETKT 360
DB 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETKT 409
QY 361 SLSFRSOMLAMIETKLNQNSRSEAPNWTQDSGFY 396
DB 410 SLSFRSOMLAMIETKLNQNSRSEAPNWTQDSGFY 445

RESULT 6
Q6IAX5 PRELIMINARY; PRT; 445 AA.
AC Q6IAX5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Eukaryotic translation initiation factor 3, subunit 6.
DE EIF3S6 protein.
GN Name=EIF3S6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR457029; CNG33310.1; -.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
SQ SEQUENCE 445 AA; 52205 MW; C7C0FF755BD9A8D4 CRC64;

Query Match          99.7%; Score 2064; DB 2; Length 445;
Best Local Similarity 99.7%; Pred. No. 1.9e-144;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDI PHALREKRTTVVAQLKQLQAEETPIVKMFEDPETTRQMOSTRD 60
DB 50 MVDFAVDVYKNLYSDDI PHALREKRTTVVAQLKQLQAEETPIVKMFEDPETTRQMOSTRD 109
QY 61 GRMLFDYLDKKGFRQYEDTLRYAKFOYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
DB 110 GRMLFDYLDKKGFRQYEDTLRYAKFOYECGNYSGAAEYLYFFRVLPATDRNALSSL 169
QY 121 WGLKASEILMQNWDAAEMDLTKETIDNNSSVSPLOSLQORTWLHWSLFFVFNHPKGR 180
DB 170 WGLKASEILMQNWDAAEMDLTKETIDNNSSVSPLOSLQORTWLHWSLFFVFNHPKGR 229
```

```
QY 181 DNIIDLFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRROVLKDLVKVIOQESYTYK 240
DB 230 DNIIDLFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRROVLKDLVKVIOQESYTYK 289
QY 241 DPITEFVECLYVNFDFDGAQKLCRECEVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB 290 DPITEFVECLYVNFDFDGAQKLCRECEVLVNDFFLVACLEDFIENARLFIETFCRIHQ 349
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETKT 360
DB 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETKT 409
QY 361 SLSFRSOMLAMIETKLNQNSRSEAPNWTQDSGFY 396
DB 410 SLSFRSOMLAMIETKLNQNSRSEAPNWTQDSGFY 445

RESULT 7
Q6P7L9 PRELIMINARY; PRT; 446 AA.
AC Q6P7L9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Eukaryotic translation initiation factor 3, subunit 6.
DE Name=eif3-p48-prov;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T.H., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madañ A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; SC061611; AAH61611.1; -.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
DR Initiation factor.
SQ SEQUENCE 446 AA; 52296 MW; EB12CF70898DA8B1 CRC64;

Query Match          96.9%; Score 2006.5; DB 2; Length 446;
Best Local Similarity 96.5%; Pred. No. 3.4e-140;
Matches 383; Conservative 9; Mismatches 4; Indels 1; Gaps 1;
```

QY 1 MVDFAVDVYKNLYSD-DIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTR 59  
DB 50 MVDFAVDVYKNLYSD-DIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTR 109  
QY 60 DGRMLFDYLDKKGFRQBYLDLYRYAKFQYECGNYSGAAEYLYFFRVLVLPATDRLNALSS 119  
DB 110 DGRMLFDHLEAKKHGFRQBYLDLYRYAKFQYECGNYSGAAEYLYFFRVLVLPSTDRNALSS 169  
QY 120 LGWKLASILMQLNWDAAEMEDTLRLKETIDNNSSVPSLOQRTWLIIHWSLVFVFNHPKG 179  
DB 170 LGWKLASILMQLNWDAAEMEDTLRLKETIDNNSSVPSLOQRTWLIIHWSLVFVFNHPKG 229  
QY 180 RDNIDFLYQPYLNAIQTCMCHLRYLTAVITNKDVRKRRQVLDKLVKVIQOESYTY 239  
DB 230 RDNIDFLYQPYLNAIQTCMCHLRYLTAVITNKDVRKRRQVLDKLVKVIQOESYTY 289  
QY 240 KDPITEFVECLYVNFDFDGAQKLRCECVSLVNDFFLVACLEDFTIENARLFIPTFCRIH 299  
DB 290 KDPITEFVECLYVNFDFDGAQKLRCECVSLVNDFFLVACLEDFTIENARLFIPTFCRIH 349  
QY 300 QCISINMLADKLNTMTPPEAEERWVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIKT 359  
DB 350 QCISINMLADKLNTMTPPEAEERWVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIKT 409  
QY 360 KSLSPRSQMLAMNIEKLNONSRSAPNWAQDSGFY 396  
DB 410 KSLAFRSQMLAMNIEKLNONSRSAPNWAQDSGFY 446

RESULT 8  
Q7ZXAS PRELIMINARY; PRT; 423 AA.  
ID Q7ZXAS  
AC Q7ZXAS  
DT 01-JUN-2003 (Tremblrel. 24, Created)  
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
DT 01-WAR-2004 (Tremblrel. 26, Last annotation update)  
DE LOC398503 protein (Fragment).  
GN Name=LOC398503;  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skala U., Smalish D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RA "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson F.;

RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative";  
RL Dev. Dyn. 225:384-391 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Strausberg R.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC045079; AAH45079.1; -;  
DR InterPro; IPR000717; PCI.  
DR InterPro; IPR010935; SMC\_hinge.  
DR Pfam; PF01399; PCI; 1.  
DR SMART; SM00088; PINT; 1.  
FT NON TER 1  
SQ SEQUENCE 423 AA; 49523 MW; F691314F43EF400B CRC64;  
Query Match 96.2%; Score 1992.5; DB 2; Length 423;  
Best Local Similarity 96.0%; Pred. No. 3.5e-139;  
Matches 381; Conservative 10; Mismatches 5; Indels 1; Gaps 1;  
QY 1 MVDFAVDVYKNLYSD-DIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTR 59  
DB 27 MVDFAVDVYKNLYSD-DIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTR 86  
QY 60 DGRMLFDYLDKKGFRQBYLDLYRYAKFQYECGNYSGAAEYLYFFRVLVLPATDRLNALSS 119  
DB 87 DGRMLFDHLEAKKHGFRQBYLDLYRYAKFQYECGNYSGAAEYLYFFRVLVLPSTDRNALSS 146  
QY 120 LGWKLASILMQLNWDAAEMEDTLRLKETIDNNSSVPSLOQRTWLIIHWSLVFVFNHPKG 179  
DB 147 LGWKLASILMQLNWDAAEMEDTLRLKETIDNNSSVPSLOQRTWLIIHWSLVFVFNHPKG 206  
QY 180 RDNIDFLYQPYLNAIQTCMCHLRYLTAVITNKDVRKRRQVLDKLVKVIQOESYTY 239  
DB 207 RDNIDFLYQPYLNAIQTCMCHLRYLTAVITNKDVRKRRQVLDKLVKVIQOESYTY 266  
QY 240 KDPITEFVECLYVNFDFDGAQKLRCECVSLVNDFFLVACLEDFTIENARLFIPTFCRIH 299  
DB 267 KDPITEFVECLYVNFDFDGAQKLRCECVSLVNDFFLVACLEDFTIENARLFIPTFCRIH 326  
QY 300 QCISINMLADKLNTMTPPEAEERWVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIKT 359  
DB 327 QCISINMLADKLNTMTPPEAEERWVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIKT 386  
QY 360 KSLSPRSQMLAMNIEKLNONSRSAPNWAQDSGFY 396  
DB 387 KSLAFRSQMLAMNIEKLNONSRSAPNWAQDSGFY 423

RESULT 9  
Q918W7 PRELIMINARY; PRT; 446 AA.  
ID Q918W7  
AC Q918W7  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-WAR-2004 (Tremblrel. 26, Last annotation update)  
DE Int-6 protein. (African clawed frog).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Morris-Desbois C., Bochar V., Reynaud C., Jalinot P.;  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF162775; AAF80474.1; -;  
DR InterPro; IPR000717; PCI.  
DR InterPro; IPR010935; SMC\_hinge.  
DR Pfam; PF01399; PCI; 1.  
DR SMART; SM00088; PINT; 1.  
SQ SEQUENCE 446 AA; 52240 MW; 21C5605DA9CF88BF CRC64;

Query Match 95.7%; Score 1981.5; DB 2; Length 446;  
 Best Local Similarity 95.2%; Pred. No. 2.4e-138;  
 Matches 378; Conservative 12; Mismatches 6; Indels 1; Gaps 1;

QY 1 MVDPMDDVYKLYSD-DIPHALREKTTTVAQKQLQAEETPIVKMFEDPETTRQMSTR 59  
 DB 50 MVDPMDDVYKLYSD-DIPHALREKTTTVAQKQLQAEETPIVKMFEDPETTRQMSTR 109

QY 60 DGRMLFDYLDKGGFQREYLDLTLYRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSS 119  
 DB 110 DGRMLFDYLDKGGFQREYLDLTLYRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSS 169

QY 120 LKGLKASEILMOWDAAMEDLTRKETIDNNSSVSPLOSLQKQRTWLHSLFVFFNHPKG 179  
 DB 170 LKGLKASEILMOWDAAMEDLTRKETIDNNSSVSPLOSLQKQRTWLHSLFVFFNHPKG 229

QY 180 RDNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLKDLVKVYQOESYTY 239  
 DB 230 RDNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLKDLVKVYQOESYTY 289

QY 240 KDPITFEVCLYVNFDFDGAOKKLRECEVSNDFLVAACLEDFIENARLFIPETFCRIH 299  
 DB 290 KDPITFEVCLYVNFDFDGAOKKLRECEVSNDFLVAACLEDFIENARLFIPETFCRIH 349

QY 300 QCSISIMLADKLNMTPEEAERWIVNLRNARLDKIDSGLGHVVMGNNVSPYQOVIK 359  
 DB 350 QCSISIMLADKLNMTPEEAERWIVNLRNARLDKIDSGLGHVVMGNNVSPYQOVIK 409

QY 360 KSLSFQSLMANNIEKKLNQNSSEAPNATQDSGFY 396  
 DB 410 KSLSFQSLMANNIEKKLNQNSSEAPNATQDSGFY 446

## RESULT 10

Q6P7X8 ID Q6P7X8 PRELIMINARY; PRT; 446 AA.  
 AC Q6P7X8;  
 DT 05-JUL-2004 (TReMBLrel. 27, Created)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
 DE Eukaryotic translation initiation factor 3, subunit 6.  
 GN ORFNames=zgc:63821;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB; TISSUE=Whole body;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Sklasko U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=AB; TISSUE=Whole body;  
 RA Strausberg R.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC061454; AAH61454.1; -;  
 DR ZFIN; ZDB-GENE-030131-3827; zgc:63821.  
 DR GO; GO:0003743; F:translation initiation factor activity; IEA.  
 DR InterPro; IPR000717; PCI.  
 DR InterPro; IPR010935; SMC\_hinge.  
 DR Pfam; PF01399; PCI; 1.  
 DR SMART; SM00088; PINT; 1.  
 KW Initiation factor.  
 SQ SEQUENCE 446 AA; 52389 MW; 32FD9D000E69F4F87 CRC64;

Query Match 94.3%; Score 1952.5; DB 2; Length 446;  
 Best Local Similarity 93.2%; Pred. No. 3.4e-136;  
 Matches 370; Conservative 18; Mismatches 8; Indels 1; Gaps 1;

QY 1 MVDPMDDVYKLYSD-DIPHALREKTTTVAQKQLQAEETPIVKMFEDPETTRQMSTR 59  
 DB 50 MVDPMDDVYKLYSD-DIPHALREKTTTVAQKQLQAEETPIVKMFEDPETTRQMSTR 109

QY 60 DGRMLFDYLDKGGFQREYLDLTLYRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSS 119  
 DB 110 DGRMLFDYLDKGGFQREYLDLTLYRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSS 169

QY 120 LKGLKASEILMOWDAAMEDLTRKETIDNNSSVSPLOSLQKQRTWLHSLFVFFNHPKG 179  
 DB 170 LKGLKASEILMOWDAAMEDLTRKETIDNNSSVSPLOSLQKQRTWLHSLFVFFNHPKG 229

QY 180 RDNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLKDLVKVYQOESYTY 239  
 DB 230 RDNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLKDLVKVYQOESYTY 289

QY 240 KDPITFEVCLYVNFDFDGAOKKLRECEVSNDFLVAACLEDFIENARLFIPETFCRIH 299  
 DB 290 KDPITFEVCLYVNFDFDGAOKKLRECEVSNDFLVAACLEDFIENARLFIPETFCRIH 349

QY 300 QCSISIMLADKLNMTPEEAERWIVNLRNARLDKIDSGLGHVVMGNNVSPYQOVIK 359  
 DB 350 QCSISIMLADKLNMTPEEAERWIVNLRNARLDKIDSGLGHVVMGNNVSPYQOVIK 409

QY 360 KSLSFQSLMANNIEKKLNQNSSEAPNATQDSGFY 396  
 DB 410 KSLSFQSLMANNIEKKLNQNSSEAPNATQDSGFY 446

RESULT 11  
 Q6DRI1 ID Q6DRI1 PRELIMINARY; PRT; 446 AA.  
 AC Q6DRI1;  
 DT 25-OCT-2004 (TReMBLrel. 28, Created)  
 DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)  
 DE Eukaryotic translation initiation factor 3 subunit 6.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=15256591; DOI=10.1073/pnas.0403929101;  
 RA Amsterdam A., Nissen R.M., Sun Z., Swindell E.C., Farrington S.,  
 RA Hopkins N.;  
 RT "Identification of 315 genes essential for early zebrafish  
 RT development."  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:12792-12797(2004).  
 DR EMBL; AY648778; AAT68096.1; -;  
 DR GO; GO:0003743; F:translation initiation factor activity; IEA.  
 DR InterPro; IPR000717; PCI.  
 DR InterPro; IPR010935; SMC\_hinge.  
 DR Pfam; PF01399; PCI; 1.  
 DR SMART; SM00088; PINT; 1.

```
KW Initiation factor.
SQ SEQUENCE 446 AA; 52398 MW; 2C36E460A8002BB7 CRC64;

Query Match 94.0%; Score 1947.5; DB 2; Length 446;
Best Local Similarity 92.9%; Pred. No. 8e-136;
Matches 369; Conservative 18; Mismatches 9; Indels 1; Gaps 1;

QY 1 MYDFAMDVYKNLYSD-DIPHALREKRTTVAOLKQLOAETEPVKMFEDPPTTROMQSTR 59
DB 50 MYDFAMDVYKNLYSD-DIPHALREKRTTVAOLKQLOAETEPVKMFEDPPTTROMQSTR 109

QY 60 DGRMLFDYLADKHGRFQRYELDTLYRYAKFOYECGNYSGAAEYLYFRVLVPAVDNRNALSS 119
DB 110 DGRMLFDYLADKHGRFQRYELDTLYRYAKFOYECGNYSGAAEYLYFRVLVPAVDNRNALSS 169

QY 120 LMGKLASETLMQNWDAAMEDLTRKKTETIDNNVSSPQSLQRTWLIIHWSLFFVFNHPKG 179
DB 170 LMGKLASETLMQNWDAAMEDLTRKKTETIDNNVSSPQSLQRTWLIIHWSLFFVFNHPKG 229

QY 180 RNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKVKRRQVLDKLVKVIQOESYTY 239
DB 230 RNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKVKRRQVLDKLVKVIQOESYTY 289

QY 240 KDPITEFEVCLYVNFDFDGAQKLRCEESVLVNDFFLVACLDFEINARLFIETFCRIH 299
DB 290 KDPITEFEVCLYVNFDFDGAQKLRCEESVLVNDFFLVACLDFEINARLFIETFCRIH 349

QY 300 QCISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIKT 359
DB 350 QCISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIKT 409

QY 360 KSLSPRSQMLANNIEKLNQNSRSEAPNATODSGFY 396
DB 410 KSLSPRSQMLANNIEKLNQNSRSEAPNATODSGFY 446

RESULT 12
Q9CT23 PRELIMINARY; PRT; 422 AA.
AC Q9CT23;
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 26, Last annotation update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:2610017H09 product:mammary tumor integration
DE site 6, full insert sequence. (fragment).
GN Name=Eif396;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=92729253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
```

```

RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Itoh M.,
RA Kanno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Sumi N., Ishii Y., Nishikawa K., Katsunai T., Tashiro H., Itoh M.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL, AK011436; BAB27621.1; -.
DR MGI, MGI:99257; Eif396.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
FT NON_TER 422
SQ SEQUENCE 422 AA; 49568 MW; A05157DBC6D375B CRC64;

Query Match 94.0%; Score 1946; DB 2; Length 422;
Best Local Similarity 100.0%; Pred. No. 9.6e-136;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDFAMDVYKNLYSD-DIPHALREKRTTVAOLKQLOAETEPVKMFEDPPTTROMQSTR 60
DB 50 MYDFAMDVYKNLYSD-DIPHALREKRTTVAOLKQLOAETEPVKMFEDPPTTROMQSTR 109

QY 61 GRMLFDYLADKHGRFQRYELDTLYRYAKFOYECGNYSGAAEYLYFRVLVPAVDNRNALSS 120
DB 110 GRMLFDYLADKHGRFQRYELDTLYRYAKFOYECGNYSGAAEYLYFRVLVPAVDNRNALSS 169

QY 121 WGLKLASETLMQNWDAAMEDLTRKKTETIDNNVSSPQSLQRTWLIIHWSLFFVFNHPKG 180
DB 170 WGLKLASETLMQNWDAAMEDLTRKKTETIDNNVSSPQSLQRTWLIIHWSLFFVFNHPKG 229

QY 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKVKRRQVLDKLVKVIQOESYTY 240
DB 230 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKVKRRQVLDKLVKVIQOESYTY 289

QY 241 DPITEFEVCLYVNFDFDGAQKLRCEESVLVNDFFLVACLDFEINARLFIETFCRIH 300
DB 290 DPITEFEVCLYVNFDFDGAQKLRCEESVLVNDFFLVACLDFEINARLFIETFCRIH 349

QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIKT 360
DB 350 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIKT 409
```

Db 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIEKTK 409  
QY 361 SLSFRSOMLANNI 373  
Db 410 SLSFRSOMLANNI 422  
RESULT 13  
Q6TH16  
ID 06TH16 PRELIMINARY; PRT; 448 AA.  
AC Q6TH16;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Eukaryotic translation initiation factor 3, subunit 6 48kDa.  
GN Names-E1F3S6; ORFNames-zgc:63821;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
[1]  
SEQUENCE FROM N.A.  
RC TISSUE=Kidney marrow;  
RA Song H.D., Wu X.Y., Sun X.J., Zhou Y., Liu T.X., Deng M., Zhang G.W.,  
RA Sheng Y., Chen Y., Ruan Z., Jiang C.L., Fan H.Y., Zou L.I.,  
RA Kanki J.P., Look A.T., Chen Z.;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY398340; AAQ97773.1; -  
DR ZFIN; ZDB-GENE-030131-3827; zgc:63821.  
DR GO; GO:0003743; F:translation initiation factor activity; IEA.  
DR InterPro; IPR00717; PCI.  
DR InterPro; IPR010935; SMC\_hinge.  
DR Pfam; PF01399; PCI; 1.  
DR SMART; SM00088; PINT; 1.  
KW Initiation factor.  
SQ SEQUENCE 448 AA; 52672 MW; 190C9FD4CEB77803 CRC64;

Query Match 86.1%; Score 1783.5; DB 2; Length 448;  
Best Local Similarity 84.6%; Pred. No. 1.1e-123;  
Matches 346; Conservative 24; Mismatches 16; Indels 23; Gaps 3;  
QY 1 MVDFAMDVKNLYSD-DIPHALREKETTVAQKQLOAETEPVKMFEPTTQMOSTR 59  
Db 50 MVDFAMDVYRNLPDPKEIPNSUREKETTVAQKQLOAETEPVKVFEPTTQMOSTR 109  
QY 60 DGRMLFDYLADKHGFRQYEDTLRYAKFOYECGNYSGAAEYVFRVLVPAIDRNALGS 119  
Db 110 DGRMLFDYLADKHGFRQYEDTLRYAKFOY-----VWELLVQRSTCTSSRS 159  
QY 120 -----LWGLASEILMQNWDAAEMEDTLRLKETIDNNSSVSPQLSQLOQTWLIH 167  
Db 160 GPCHRQKRAELLWGLASEILMQNWEAEMEDTLRLRETIDNNVSSVSPQLSQLOQTWLIH 219  
QY 168 WSLFVFFNPKGRDNIIDFLQPOVYLAQIOTMCPHILRYLTAVITNKKVRQVLRKD 227  
Db 220 WSLFVFFNPKGRDNIIEFLQPOVYLAQIOTMCPHILRYLTAVITNKKVRQVLRKD 279  
QY 228 LVKVIQESYTYKDPITEFVECLVYNDFDGAQKLRCECVLNDFFLVACLEDPIENA 287  
Db 280 LVKVIQESYTYKDPITEFVECLVYNDFDGAQKLRCECVLNDFFLVACLEDPIENA 339  
QY 288 RLFIETFCRIHQICISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGN 347  
Db 340 RLFIETFCRIHQICISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGN 399  
QY 348 AVSPYQOVIEKTKSLSFRSOMLANNIEKKLNNSREAPNATQSGFY 396  
Db 400 AISPYQOVIEKTKSLSFRSOMLANNIEKKLNNSREAPNATQSGFY 448

RESULT 14  
IF36\_DROME

ID IF36\_DROME STANDARD; PRT; 435 AA.  
AC Q774I0; Q9VVA2;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 25-JAN-2005 (Rel. 46, Last annotation update)  
DE Probable eukaryotic translation initiation factor 3 subunit 6 (eIF-3 p48) (eIF3e).  
GN Name-Int6; ORFNames-CG9677;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
[1]  
SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Miyazaki S., Diella F., Callahan D., Sullivan D., Callahan R.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
[2]  
SEQUENCE FROM N.A.  
RC STRAIN=Berkely; TISSUE=Ovary;  
RX MEDLINE=20196012; PubMed=10731138; DOI=10.1126/science.287.5461.2222;  
RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Frise E.,  
RA Stapleton M., Harvey D.A.;  
RT "A Drosophila complementary DNA resource.";  
RL Science 287:2222-2224 (2000).  
[3]  
SEQUENCE FROM N.A.  
RC STRAIN=Berkely;  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhargava S., Boldakovic S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin C.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fouts C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195 (2000).  
[4]  
GENOME REANNOTATION.  
RP MEDLINE=22426069; PubMed=12537572;  
RX Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bertencourt B.R., Celiniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 CC -!- FUNCTION: Binds to the 40S ribosome and promotes the binding of  
 CC methionyl-tRNAi and mRNA (by similarity).  
 CC -!- SUBUNIT: eIF-3 is composed of at least 12 different subunits (by  
 CC similarity).  
 CC -!- SIMILARITY: Belongs to the eIF3S6 family.  
 CC -!- SIMILARITY: Contains 1 PCI domain.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U9162; AAC62307.1; -.  
 DR EMBL; AF132551; AAD27850.1; -.  
 DR EMBL; AE003526; AAF49412.1; -.  
 DR FlyBase; FBgn0025582; Int6.  
 DR InterPro; IPR000717; PCI.  
 DR Pfam; PF01399; PCI; 1.  
 DR SMART; SM00088; PINT; 1.  
 KW Initiation factor; Protein biosynthesis.  
 SQ SEQUENCE 435 AA; 51162 MW; 6F79D32FBA3EA711 CRC64;

Query Match 58.9%; Score 1220.5; DB 1; Length 435;  
 Best Local Similarity 59.7%; Pred. No. 4.9e-82;  
 Matches 233; Conservative 65; Mismatches 83; Indels 9; Gaps 5;  
 QY 1 MVDFAMDVYKNI-YSDDIPHALREKRTTVVAOLKQLAETEPVIVKMFDPETTRQVQSTR 59  
 DB 50 MIDYTMTRKRLNLSQEMPEELVQKAEVLAATLQKQNEVAPIMKA---TDILKNGESWK 106  
 QY 60 DGRMLFDYLDKHFGRQBYLDLTLYRYAKFOYECGNYSGAAEYLYFFRVLPVATDRNALSS 119  
 DB 107 DSKTFNALQKYNFKVEHLESAKLYKLYECGNYQESTSYLYFCLIVMSPNCKNLYNV 166  
 QY 120 LRGKLAELILNWNNTALEDLTLRDLTDYIDNANFST-IQALQQRQTLIHWSVLVFFNHKPG 179  
 DB 167 LAGKLAELILNWNNTALEDLTLRDLTDYIDNANFST-IQALQQRQTLIHWSVLVFFNHKPG 225  
 QY 180 RDNIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIQOESYTY 239  
 DB 226 RDLIIEMFLYKPLYLNAIQTMCPHIMRYLATAVIN---RTRNALKDLIKVIQOESYTY 282  
 QY 240 KDPITEFVCLVYNFDFGAOKKLRECESVLVNDFLVACLEDFIENARLFIETFCRIH 299  
 DB 283 RDPITEFVCLVYNFDFGAOKKLRECESVLVNDFLVACLEDFIENARLFIETFCRIH 342  
 QY 300 QCISINMLADKLMNTPPEEAERWIVNLIRNARLDKIDS KLGHVVMGNNAVSPYQVIEKT 359  
 DB 343 QCITISMLADKLMNTPPEEAERWIVNLIRNARLDKIDS KLGHVVMGNNAVSPYQVIEKT 402  
 QY 360 KSLSFQSMQMLAMNIEKLNQNSRSE-APNW 388  
 DB 403 DLSLSRSEHLAGLIERKSKQKQNGESADSW 432

## RESULT 15

Q7QIL8 PRELIMINARY; PRT; 462 AA.  
 AC Q7QIL8;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE AGCP3626 (Fragment).  
 GN Name=agCG57001; ORFNames=ENSAANG00000019251;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota,  
 OC Neoptera, Endopterygota, Diptera, Nematocera, Culicoidea, Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAB01008807; EAA04727.1; -.  
 DR InterPro; IPR000717; PCI.  
 DR Pfam; PF01399; PCI; 1.  
 FT NON TER 1  
 SQ SEQUENCE 462 AA; 53647 MW; 835173A6776AA4E0 CRC64;  
 Query Match 56.3%; Score 1165.5; DB 2; Length 462;  
 Best Local Similarity 58.0%; Pred. No. 6.3e-78;  
 Matches 232; Conservative 56; Mismatches 83; Indels 29; Gaps 7;  
 QY 1 MVDFAMDVYKNIYSDDI-PHALREKRTTVVAOLKQLAETEPVIVKMFDPETTRQVQSTR 59  
 DB 79 MVDYKHDIRERLAMDKVHPDELAQGRANVLATLKELOAEVAPLMKMBE---LKNPDSK 135  
 QY 60 DGRMLFDYLDKHFGRQBYLD-----TLRYAKFOYECGNYSGAAEYLYFFRVLPVATDR 114  
 DB 136 DSKSVIHAI-----QOTLDYDIILSAQKLAKLYECGNYNDSLYLVCMVMEPNDK 188  
 QY 115 NALSSLGKGLASEILMQNWDAAEDTLRLKETIDNNSVSSPLQSQOQRTWLHWSLVFVF 174  
 DB 189 NYLGVWLGKLAVEIITLWQTALEDLTLRDLTDYIDNANFST-IQALQQRQTLIHWSVLVFF 247  
 QY 175 NHPKGRDNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIQO 234  
 DB 248 NHGKGRDLIIDMFLYKPOYLNAIQTMCPHILRYLATAVIN---RGRNALKDLIKVIQO 304  
 QY 235 ESYTYKDPITEFVCLVYNFDFGAOKKLRECESVLVNDFLVACLEDFIENARLFIET 294  
 DB 305 ESYTYRDPITEFVCLVYNFDFGAOKKLRECESVLVNDFLVACLEDFIENARLFIET 364  
 QY 295 FCRHQCISINMLADKLMNTPPEEAERWIVNLIRNARLDKIDS KLGHVVMGNNAVSPYQO 354  
 DB 365 FCRHQCITIGMLADKLMNTPPEEAERWIVNLIRNARLDKIDS KLGHVVMGNNAVSPYQO 424  
 QY 355 VIEKTSKLSFRSQMLAMNIEKLNQNSRSEAPNWTQDSG 394  
 DB 425 LIVEKIDSLSVRSEALTLLVERKHAK-----TQEAG 455

Search completed: March 8, 2005, 20:07:17  
 Job time : 175 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 20:07:23 ; Search time 136 Seconds  
(without alignments)  
958.101 Million cell updates/sec

Title: US-10-783-415-4  
Perfect score: 2071  
Sequence: 1 MVDAMDVYKNLYSDIPHA.....LQNSRSEAPNWTQDSGFY 396

Scoring table: BLOSUM62

Searched: 1391452 seqs, 32904822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pcp:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pcp:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pcp:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pcp:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pcp:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pcp:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pcp:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pcp:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pcp:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pcp:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pcp:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pcp:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pcp:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pcp:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pcp:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pcp:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pcp:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pcp:\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pcp:\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2071	100.0	396	10	US-09-858-152A-2
2	2071	100.0	396	10	US-09-858-152A-4
3	2071	100.0	396	16	US-10-408-765A-2064
4	2071	100.0	396	16	US-10-783-415-2
5	2071	100.0	396	16	US-10-783-415-4
6	2071	100.0	445	13	US-10-087-192-558
7	1279.5	61.8	320	13	US-10-087-192-558
8	1087.5	52.5	439	16	US-10-767-701-45301
9	1085.5	52.4	439	16	US-10-437-963-188307
10	1064	51.4	437	15	US-10-424-599-266298
11	952.5	46.0	444	16	US-10-437-963-188308
12	564	27.2	117	15	US-10-264-049-4238
13	555	26.8	187	15	US-10-425-114-61814

14	391.5	18.9	130	15	US-10-424-599-217317	Sequence 217317,
15	260	12.6	166	16	US-10-437-963-188446	Sequence 188446,
16	223	10.8	45	9	US-09-864-761-42338	Sequence 42338, A
17	211	10.2	42	9	US-09-864-761-40729	Sequence 40729, A
18	211	10.2	42	9	US-09-864-761-42138	Sequence 42138, A
19	173	8.4	112	14	US-10-106-698-6920	Sequence 6920, Ap
20	135	6.5	61	15	US-10-424-599-213418	Sequence 213418,
21	114	5.5	381	15	US-10-424-599-225049	Sequence 225049,
22	112	5.4	567	13	US-10-087-192-1170	Sequence 1170, Ap
23	110	5.3	858	15	US-10-425-114-72451	Sequence 72451, A
24	108	5.2	147	15	US-10-425-114-49583	Sequence 49583, A
25	107.5	5.2	641	16	US-10-444-693C-36	Sequence 36, Appl
26	106.5	5.1	1186	15	US-10-282-122A-60346	Sequence 60346, A
27	106	5.1	1146	15	US-10-282-122A-66407	Sequence 66407, A
28	106	5.1	1575	16	US-10-408-765A-704	Sequence 704, App
29	105	5.1	1323	15	US-10-389-566-1642	Sequence 1642, Ap
30	104.5	5.0	681	15	US-10-282-122A-71657	Sequence 71657, A
31	104.5	5.0	684	15	US-10-104-047-2418	Sequence 2418, Ap
32	103.5	5.0	618	14	US-10-203-860-24	Sequence 24, Appl
33	103.5	5.0	742	14	US-10-203-860-2	Sequence 2, Appli
34	103.5	5.0	742	14	US-10-308-448-11	Sequence 11, Appl
35	103.5	5.0	742	15	US-10-341-434-85	Sequence 85, Appl
36	102.5	4.9	806	16	US-10-437-963-149528	Sequence 149528,
37	101.5	4.9	872	14	US-10-326-040-7	Sequence 7, Appli
38	101	4.9	1039	15	US-10-369-493-16475	Sequence 16475, A
39	100.5	4.9	87	15	US-10-424-599-168200	Sequence 168200,
40	100.5	4.9	872	15	US-10-424-599-219651	Sequence 219651,
41	100.5	4.9	1020	15	US-10-282-122A-58016	Sequence 58016, A
42	100	4.8	1575	17	US-10-741-600-1548	Sequence 1548, Ap
43	100	4.8	1575	17	US-10-741-600-1549	Sequence 1549, Ap
44	100	4.8	1640	17	US-10-741-600-1550	Sequence 1550, Ap
45	99.5	4.8	4563	16	US-10-741-601-432	Sequence 432, App

ALIGNMENTS

RESULT 1  
US-09-858-152A-2  
; Sequence 2, Application US/09858152A  
; Publication No. US2003004419A1  
; GENERAL INFORMATION:  
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE  
; APPLICANT: SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES  
; APPLICANT: Marchetti, Antonio  
; APPLICANT: Buttitta, Flamma  
; APPLICANT: Smith, Gilbert H.  
; APPLICANT: Callahan, Robert  
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6  
; FILE REFERENCE: 4239-59122  
; CURRENT APPLICATION NUMBER: US/09/858,152A  
; CURRENT FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Murine INT6  
US-09-858-152A-2

Query Match 100.0%; Score 2071; DB 10; Length 396;  
Best Local Similarity 100.0%; Pred. No. 7.2e-179; Indels 0; Gaps 0;  
Matches 396; Conservative 0; Mismatches 0;  
QY 1 MVDAMDVYKNLYSDIPHALEKRTTVVAQKQLQAEPIVKMFEDPETTRQMSTRD 60  
Db 1 MVDAMDVYKNLYSDIPHALEKRTTVVAQKQLQAEPIVKMFEDPETTRQMSTRD 60  
QY 61 GRMLFDYLDKGFQREYLDTLRYAKFOYECNYSGAAYLYFFRVLPATDRNALSSL 120  
Db 61 GRMLFDYLDKGFQREYLDTLRYAKFOYECNYSGAAYLYFFRVLPATDRNALSSL 120  
QY 121 WGLASEILMQNWDAAEDLTRLKETIDNNSVSSQLQORTWLHWSLFVFFNPKGR 180

```
Db 121 WGLKASEILMQNWDAAEDLTRELKETIDNNSVSSPLQSLQRTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPOLNAINAQTMCPIHLRYLTAVITNKDVRKRQVLKDLVKVLIQOESYTK 240
Db 181 DNIIDFLYQPOLNAINAQTMCPIHLRYLTAVITNKDVRKRQVLKDLVKVLIQOESYTK 240
QY 241 DPITEFVECLYVNFDFDGAQKKLRECVSLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKKLRECVSLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWTINLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEEAERWTINLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
QY 361 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 2
US-09-858-152A-4
; Sequence 4, Application US/09858152A
; Publication No. US2003004419A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE
; APPLICANT: SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Fianna
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
; CURRENT APPLICATION NUMBER: US/09/858,152A
; CURRENT FILING DATE: 2001-05-14
; SOFTWARE: PatentIn version 3.1
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 4
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-152A-4

Query Match 100.0%; Score 2071; DB 10; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-179;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVVAQLKQLQAETEPVVKMFDPETTRMQSTRD 60
Db 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVVAQLKQLQAETEPVVKMFDPETTRMQSTRD 60
QY 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
Db 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
QY 121 WGLKASEILMQNWDAAEDLTRELKETIDNNSVSSPLQSLQRTWLIHWSLFFVFNHPKGR 180
Db 121 WGLKASEILMQNWDAAEDLTRELKETIDNNSVSSPLQSLQRTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPOLNAINAQTMCPIHLRYLTAVITNKDVRKRQVLKDLVKVLIQOESYTK 240
Db 181 DNIIDFLYQPOLNAINAQTMCPIHLRYLTAVITNKDVRKRQVLKDLVKVLIQOESYTK 240
QY 241 DPITEFVECLYVNFDFDGAQKKLRECVSLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKKLRECVSLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWTINLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEEAERWTINLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
QY 361 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

US-09-858-152A-4
```

```
Db 361 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 3
US-10-408-765A-2064
; Sequence 2064, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Faby, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2064
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2064

Query Match 100.0%; Score 2071; DB 16; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-179;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVVAQLKQLQAETEPVVKMFDPETTRMQSTRD 60
Db 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVVAQLKQLQAETEPVVKMFDPETTRMQSTRD 60
QY 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
Db 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
QY 121 WGLKASEILMQNWDAAEDLTRELKETIDNNSVSSPLQSLQRTWLIHWSLFFVFNHPKGR 180
Db 121 WGLKASEILMQNWDAAEDLTRELKETIDNNSVSSPLQSLQRTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPOLNAINAQTMCPIHLRYLTAVITNKDVRKRQVLKDLVKVLIQOESYTK 240
Db 181 DNIIDFLYQPOLNAINAQTMCPIHLRYLTAVITNKDVRKRQVLKDLVKVLIQOESYTK 240
QY 241 DPITEFVECLYVNFDFDGAQKKLRECVSLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKKLRECVSLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWTINLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEEAERWTINLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
QY 361 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

US-10-783-415-2
; Sequence 2, Application US/10783415
; Publication No. US20040141918A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Fianna
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
```



;; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6  
;; FILE REFERENCE: 4239-59122  
;; CURRENT APPLICATION NUMBER: US/10/783,415  
;; CURRENT FILING DATE: 2004-02-19  
;; PRIOR APPLICATION NUMBER: 09/858,152  
;; PRIOR FILING DATE: 2001-05-14  
;; NUMBER OF SEQ ID NOS: 36  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 2  
;; LENGTH: 396  
;; TYPE: PRT  
;; ORGANISM: Murine INT6  
US-10-783-415-2

Query Match 100.0%; Score 2071; DB 16; Length 396;  
Best Local Similarity 100.0%; Pred. No. 7.2e-179;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MVDPMVVKYKLYSDDI PHALREKRTTVVAQLKQLQAEETPIVKMFEDPETTRQMOSTRD 60  
DB 1 MVDPMVVKYKLYSDDI PHALREKRTTVVAQLKQLQAEETPIVKMFEDPETTRQMOSTRD 60  
  
QY 61 GRMLFDYLDKKGFRQYLDTLRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120  
DB 61 GRMLFDYLDKKGFRQYLDTLRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120  
  
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 180  
DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 180  
  
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240  
DB 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240  
  
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300  
DB 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300  
  
QY 301 CISINMLADKLANTPPEAERWIINLRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTK 360  
DB 301 CISINMLADKLANTPPEAERWIINLRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTK 360  
  
QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396  
DB 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

## RESULT 5

US-10-783-415-4  
;; Sequence 4, Application US/10783415  
;; Publication No. US20040141918A1  
;; GENERAL INFORMATION:  
;; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS  
;; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND  
;; APPLICANT: HUMAN SERVICES  
;; APPLICANT: Marchetti, Antonio  
;; APPLICANT: Buttitta, Flamma  
;; APPLICANT: Smith, Gilbert H.  
;; APPLICANT: Callahan, Robert  
;; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6  
;; FILE REFERENCE: 4239-59122  
;; CURRENT APPLICATION NUMBER: US/10/783,415  
;; CURRENT FILING DATE: 2004-02-19  
;; PRIOR APPLICATION NUMBER: 09/858,152  
;; PRIOR FILING DATE: 2001-05-14  
;; NUMBER OF SEQ ID NOS: 36  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 4  
;; LENGTH: 396  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-783-415-4

Query Match 100.0%; Score 2071; DB 16; Length 396;  
Best Local Similarity 100.0%; Pred. No. 7.2e-179;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MVDPMVVKYKLYSDDI PHALREKRTTVVAQLKQLQAEETPIVKMFEDPETTRQMOSTRD 60  
DB 1 MVDPMVVKYKLYSDDI PHALREKRTTVVAQLKQLQAEETPIVKMFEDPETTRQMOSTRD 60  
  
QY 61 GRMLFDYLDKKGFRQYLDTLRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120  
DB 61 GRMLFDYLDKKGFRQYLDTLRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120  
  
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 180  
DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 180  
  
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240  
DB 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240  
  
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300  
DB 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300  
  
QY 301 CISINMLADKLANTPPEAERWIINLRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTK 360  
DB 301 CISINMLADKLANTPPEAERWIINLRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTK 360  
  
QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396  
DB 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

## RESULT 6

US-10-087-192-558  
;; Sequence 558, Application US/10087192  
;; Publication No. US20020182586A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Morris, David W.  
;; APPLICANT: Engelhard, Eric K.  
;; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
;; TITLE OF INVENTION: CANCER  
;; FILE REFERENCE: 529452000122  
;; CURRENT APPLICATION NUMBER: US/10/087,192  
;; CURRENT FILING DATE: 2002-03-01  
;; PRIOR APPLICATION NUMBER: US 09/747,377  
;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: US 09/798,586  
;; PRIOR FILING DATE: 2001-03-02  
;; NUMBER OF SEQ ID NOS: 2059  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 558  
;; LENGTH: 445  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-087-192-558

Query Match 100.0%; Score 2071; DB 13; Length 445;  
Best Local Similarity 100.0%; Pred. No. 8.5e-179;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MVDPMVVKYKLYSDDI PHALREKRTTVVAQLKQLQAEETPIVKMFEDPETTRQMOSTRD 60  
DB 50 MVDPMVVKYKLYSDDI PHALREKRTTVVAQLKQLQAEETPIVKMFEDPETTRQMOSTRD 109  
  
QY 61 GRMLFDYLDKKGFRQYLDTLRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120  
DB 110 GRMLFDYLDKKGFRQYLDTLRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 169  
  
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 180  
DB 170 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 229

QY 181 DNIIDFLYQPOLNAIOTMCPHILRYLTAVITNKDKRRQVLKDLVKVIOQESYTYK 240  
Db 230 DNIIDFLYQPOLNAIOTMCPHILRYLTAVITNKDKRRQVLKDLVKVIOQESYTYK 289  
QY 241 DPITFEVCLYVNFDFGAKKRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300  
Db 290 DPITFEVCLYVNFDFGAKKRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 349  
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNVSPYQOVIETK 360  
Db 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNVSPYQOVIETK 409  
QY 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396  
Db 410 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445

## RESULT 7

US-10-087-192-555  
; Sequence 555, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 555  
; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-087-192-555

Query Match 61.8%; Score 1279.5; DB 13; Length 320;

Best Local Similarity 66.4%; Pred. No. 2.6e-107;

Matches 263; Conservative 0; Mismatches 0; Indels 133; Gaps 1;

QY 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKOLQAETEPVVKMPEDPETTRQMQSTRD 60  
Db 58 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKOLQAETEPVVKMPEDPETTRQMQSTR- 116  
QY 61 GRMLFDYLDKRGFRQYLDTLRYAKFOYECGNYSAGAEYLFFRVLPATDRNALSSL 120  
Db 117 ----- 116  
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSPLOSLQORTWLIIHWSLFFVFNHPKGR 180  
Db 117 ----- 116  
QY 181 DNIIDFLYQPOLNAIOTMCPHILRYLTAVITNKDKRRQVLKDLVKVIOQESYTYK 240  
Db 117 -----YLNAIOTMCPHILRYLTAVITNKDKRRQVLKDLVKVIOQESYTYK 164  
QY 241 DPITFEVCLYVNFDFGAKKRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300  
Db 165 DPITFEVCLYVNFDFGAKKRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 224  
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNVSPYQOVIETK 360  
Db 225 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNVSPYQOVIETK 284  
QY 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396  
Db 285 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 320

## RESULT 8

US-10-767-701-45301  
; Sequence 45301, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5335)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 45301  
; LENGTH: 439  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C1020\_1.pap  
US-10-767-701-45301

Query Match 52.5%; Score 1087.5; DB 16; Length 439;

Best Local Similarity 53.2%; Pred. No. 1e-89;

Matches 205; Conservative 84; Mismatches 77; Indels 19; Gaps 6;

QY 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKOLQAETEPVVKMPEDPETTRQMQSTR 59  
Db 50 MVDYAMDIIHSLHGTHDDVPDMVKRAEVSRLSRLEEAAPLVAFQNPQLVQBELRP-- 107  
QY 60 DGRMLFDYLDKRGFRQYLDTLRYAKFOYECGNYSAGAEYLFFRVLPATDRNALSS 119  
Db 108 DKQYNIHMLQERYQIGPQIEALYQAKFOECGNYSAGADLYQYRALTNSER-SVSA 166  
QY 120 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSPLOSLQORTWLIIHWSLFFVFNHPK 179  
Db 167 WGLKASEILMQNWDVALEELNRLKEIIDSKNFSPLQNLRIWLMHWALSIFFNHENG 226  
QY 180 RDNIIIDFLYQPOLNAIOTMCPHILRYLTAVITNKDKRRQVLKDLVKVIOQESYTY 239  
Db 227 RNGIIDL7-FQDRLYNALQTNNAHLRLYLATAVVVYK---RRNNMLKELIKVIOQEQHTY 282  
QY 240 KDPITFEVCLYVNFDFGAKKRCESVLVNDFFL-----VACLEDFTENAR 288  
Db 283 KDPITFEVCLYVNFDFGAKKRCESVLVNDFFLGRIBEGNFVTVPLRDEFLENAR 342  
QY 289 LFIFETFCRIHQICISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNN 348  
Db 343 LFIFETFCRIHRCIDISMLSQKLNMSYDEAEELWIMNLVRSSKLDARIDSVSGTLIMTTNH 402  
QY 349 VSPYQOVIETKSLSPRSQMLAMNI 373  
Db 403 VNVHEQIIENLGLNMRTYMLAKNI 427

## RESULT 9

US-10-437-963-188307  
; Sequence 188307, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963

```
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188307
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84924C.1.pep
US-10-437-963-188307

Query Match          52.4%; Score 1085.5; DB 16; Length 439;
Best Local Similarity 54.3%; Pred. No. 1.6e-89;
Matches 209; Conservative 78; Mismatches 79; Indels 19; Gaps 6;

QY      1 MVDFAVDVKNLY-SDDIPHALREKTTTVAQLKQLQAEPTPIVKMFEDPETTRQMSTR 59
DB      50 MVDYANDIHKSLDTEVDVDDVRRRTDVSRLRALDEATPIVSFLQQLVQELRP-- 107

QY      60 DGRMLFDYLADKHGFQREYDLTLRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSS 119
DB      108 DKQYNLMLQDRFOIGPEIETUYQAKQFDCGNTSDAAVLYQYRALCTNSER-SLSA 166

QY      120 LMGKLAASEILMQNDWAAMEDLTRKETIDNNSVSSPLOSLOQRTWLIIHWSLFFVFNHPKG 179
DB      167 LMGKLAASEILMQNDWIDELNRLKEIIDSXKSPASPLNQVQSRILWMHWSLFIFFNHNG 226

QY      180 RDNIIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLDKLVKVIQESYTY 239
DB      227 RGIIDLF-FQDRIYLAQTNAPHLRLYLATAVVNKK--RRRNMLKELIKVIQEQHSY 282

QY      240 KDPITEFVECLYVNFDFDGAQKLRCEESVLVNDPFL-----VACLEDFTENAR 288
DB      283 KDPITEFVECLYVNFDFDGAQKLRCEESVLVNDPFLGRVVEESNFTVPLRDEFTENAR 342

QY      289 LFIFETFCRIHOCISINMLADKLNTPEEAERWIVNLIRNARDKIDSKLGHVVMGNNA 348
DB      343 LFIFETFCRIHOCIDIGMLSQKLSYDEAELWIMNLVNSKLDKIDSVGTLINTNH 402

QY      349 VSPYQOVIETKSLSPRSOHLNWI 373
DB      403 VNIHQVIESLKNLNRMTFLAKNI 427

RESULT 10
US-10-424-599-266298
; Sequence 266298, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 266298
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_82489C.1.pep
US-10-424-599-266298

Query Match          51.4%; Score 1064; DB 15; Length 437;
Best Local Similarity 52.4%; Pred. No. 1.4e-87;
Matches 207; Conservative 82; Mismatches 86; Indels 20; Gaps 7;

QY      1 MVDFAVDVKNLY-SDDIPHALREKTTTVAQLKQLQAEPTPIVKMFEDPETTRQMSTR 59
DB      50 MVDYANDIHKSLYHTEDVNDVDRRAEVVARLKSLEEAAPFLVAFQNEAAVQELRA-- 107
```

```
QY      60 DGRMLFDYLADKHGFQREYDLTLRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSS 119
DB      108 DKQYNLMLQDRFOIGPEIETUYQAKQFDCGNTSGAADYLYQYRALCTNSER-SLSA 166

QY      120 LMGKLAASEILMQNDWAAMEDLTRKETIDNNSVSSPLOSLOQRTWLIIHWSLFFVFNHPKG 179
DB      167 LMGKLAASEILMQNDWIDELNRLKEIIDSXKSPASPLNQVQSRILWMHWSLFIFFNHNG 226

QY      180 RDNIIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLDKLVKVIQESYTY 239
DB      227 RTQIIDLF-NQDKYLNAIQTSAPHLRLYLATAPIVKK--RRRPQKDFIKVIQEQHSY 282

QY      240 KDPITEFVECLYVNFDFDGAQKLRCEESVLVNDPFL-----VACLEDFTENAR 288
DB      283 KDPITEFVECLYVNFDFDGAQKLRCEESVLVNDPFLGRVVEESNFTVPLRDEFTENAR 342

QY      289 LFIFETFCRIHOCISINMLADKLNTPEEAERWIVNLIRNARDKIDSKLGHVVMGNNA 348
DB      343 LFIFETFCRIHOCIDIGMLSQKLSYDEAELWIMNLVNSKLDKIDSVGTLINTNH 402

QY      349 VSPYQOVIETKSLSPRSOHLNWI 382
DB      403 LNVYEQLIDHTKALNGRTYKLVSQLLEHAQAQOTTR 437

RESULT 11
US-10-437-963-188308
; Sequence 188308, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188308
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84925C.1.pep
US-10-437-963-188308

Query Match          46.0%; Score 952.5; DB 16; Length 444;
Best Local Similarity 49.9%; Pred. No. 1.8e-77;
Matches 197; Conservative 68; Mismatches 93; Indels 37; Gaps 8;

QY      1 MVDFAVDVKNLYSD-----DIPHALREKTTTVAQLKQLQAEPTPIVKMFED----P 49
DB      53 MVDYANDHTKLHGTDGDDVVVPVDDMVRRHEVTRLGALAAPAIVSALKNHHLGP 112

QY      50 ETTRQMSTRDGRMLFDYLADKHGFQREYDLTLRYAKFOYECGNTSGAAEYLYFFRVLV 109
DB      113 DKEHNI-----RMLH---ERFQIGPDQIEALYQAKFQFCGNYPDAENLHRYALC 162

QY      110 PATDRNALSSLWGLASLIMONWDAAMEDLTRKETIDNNSVSSPLOSLOQRTWLIIHWS 169
DB      163 TSSER-SUSAQMGKLSAEILNNWDAEELNCLKEMIDSKSSSSPFLNQLNRILMHCS 221

QY      170 LFVFNHPKGRDNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLDKLV 229
DB      222 IFIFNHGNGSGIIDLFPQDKYLNAIQTDAPHLRLYLAAAIIVN--RRRNMLKELV 277
```

```
QY 230 KVIQESYTKDPITEFVECLVNFDFGAOKKLECEBSVLVNDPFL-----VA 278
Db 278 KVIQEQHSYKDPITEFVECLVNFDFGAQKLECEBSVLVNDPFLGKRKEGNSITVP 337
QY 279 CLEDFIENARLFIPTFCRIHQCSINMLADKLNMTPEAEERWVNLIRNARLDKIDSK 338
Db 338 LRDFELENARLFIPTFCRIHRSIDIGMLSQKLNMRVDEGELWVNLVRSKLDKIDSV 397
QY 339 LGHVGMGNVNSPYQOVIEKTKSLFSRQMLAMNI 373
Db 398 SGTLMTTNHVNIHQFIESLKNLNMRTSMLAKNI 432

RESULT 12
US-10-264-049-4238
; Sequence 4238, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 4238
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-4238

Query Match 27.2%; Score 564; DB 15; Length 117;
Best Local Similarity 95.6%; Pred. No. 4,1e-43;
Matches 108; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 122 GKXSEILMOWNDAMEDLTRKETIDNNSVSSPLQSLQORTWLIHWSLVFVFNHPKGRD 181
Db 1 GKXSEILMOWNDAMEDLTRKETIDNNSVSSPLQSLQORTWLIHWSLVFVFNHPKGRD 60

QY 182 NIIDLFLYQPOLNAIQTMCPIHLRYLTAVITNKDVRKRRQVLKDLVKVIOQ 234
Db 61 NIIDLFLYQPOLNAIQTMCPIHLRYLTAVITNKDVRKRRQVLKDLGKVFNR 113

RESULT 13
US-10-425-114-61814
; Sequence 61814, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61814
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Zea mays
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73157G05_FLI.pep
US-10-425-114-61814

Query Match 26.8%; Score 555; DB 15; Length 187;
Best Local Similarity 57.3%; Pred. No. 5.2e-42;
Matches 102; Conservative 33; Mismatches 25; Indels 18; Gaps 3;

QY 77 EYLDTRYAKFQYECGNYSGAAEYLVFFRVLPATDRNALSSILWGKLASEILMONWDA 136
Db 23 DOIETALYQYAFQYECGNYSGAADYLYOYRALTCTNSER-VSALWGKLAABILMONWDLA 81

QY 137 MEDTLRKETIDNNSVSSPLQSLQORTWLIHWSLVFVFNHPKGRDNIIDLFLYQPOLN 196
Db 82 LEELNRLKEIIDSKNFSSPLNQLNRIWLMHWALFIPFNHIR-----YLNA 127

QY 197 IQTWCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTKDPITEFVECLVNF 254
Db 128 IQTNAHLIRYLATAVVVVK---RRNMLKELIKVIOQEHSHYKDPVTFECLVNY 182

RESULT 14
US-10-424-599-217317
; Sequence 217317, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 217317
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(130)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_38264C.1.pep
US-10-424-599-217317

Query Match 18.9%; Score 391.5; DB 15; Length 130;
Best Local Similarity 64.2%; Pred. No. 2e-27;
Matches 70; Conservative 19; Mismatches 19; Indels 1; Gaps 1;

QY 79 LDTLYRYAKFQYECGNYSGAAEYLVFFRVLPATDRNALSSILWGKLASEILMONWDA 138
Db 12 IEALYQYAKFQYECGNYSGAADYLYOYRALTCTNSER-SLSALWGKLAABILMONWDLA 70

QY 139 DLTLRKETIDNNSVSSPLQSLQORTWLIHWSLVFVFNHPKGRDNIIDLFL 187
Db 71 ELNRLKEIIDSKNFASPLNQVQSRIWLMHWSLFIFFNHDNGRTXIIDL 119

RESULT 15
US-10-437-963-188446
; Sequence 188446, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
```

Search completed: March 8, 2005, 20:19:52  
Job time : 138 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 19:49:22 ; Search time 41 Seconds  
(without alignments)  
929.313 Million cell updates/sec

Title: US-10-783-415-4  
Perfect score: 2071  
Sequence: 1 MVDFAVDYKKNLYSDDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2071	100.0	396	2 I75615	mammary tumor inte
2	994.5	48.0	432	2 T33118	hypothetical prote
3	903	43.6	418	2 T45807	translation initia
4	775	37.4	501	2 T40585	translation initit
5	760	36.7	501	2 T50488	translation initit
6	116	5.6	458	2 S56816	GTPase-activating
7	114	5.5	1158	2 T43519	condensin complex
8	114	5.5	2121	2 T27406	hypothetical prote
9	109.5	5.3	446	2 H97754	lipid-A-disacchari
10	108	5.2	481	2 G97089	gamma-glutamylcyst
11	106.5	5.1	1186	2 AD1300	Smc protein essent
12	106	5.1	1146	2 C83304	hypothetical prote
13	105.5	5.1	1186	2 AD1672	Smc protein essent
14	105	5.1	668	2 S64123	hypothetical prote
15	103.5	5.0	742	2 JC7595	scavenger receptor
16	103.5	5.0	1173	2 T30308	reXa protein - Lac
17	102	4.9	2346	2 T13829	Tpr homolog - frui
18	101.5	4.9	622	2 D84493	probable retroelme
19	100	4.8	569	1 S62851	hypothetical prote
20	99.5	4.8	300	2 T20627	hypothetical prote
21	99.5	4.8	341	2 T20626	hypothetical prote
22	99.5	4.8	732	2 T32757	hypothetical prote
23	99	4.8	519	2 T27880	hypothetical prote
24	99	4.8	1060	2 F88710	protein COI5.4 [i
25	99	4.8	1070	2 G84982	exodeoxyribonuclea
26	99	4.8	1079	2 T30996	hypothetical prote
27	99	4.8	1927	2 T25604	hypothetical prote
28	99	4.8	2541	2 T29340	hypothetical prote
29	98.5	4.8	1367	2 H82874	conserved hypothet

30	98.5	4.8	2563	1 S28261	centromere protein
31	98.5	4.8	4563	1 LPH0B	apolipoprotein B-1
32	98	4.7	760	2 T28224	ORF MSV063 probabl
33	98	4.7	979	2 JQ0894	P115 protein - Myc
34	97.5	4.7	509	2 B90573	hypothetical prote
35	97.5	4.7	761	2 T03719	probable thyroid x
36	97.5	4.7	1156	2 E69444	chromosome segrega
37	97.5	4.7	1792	2 T20363	hypothetical prote
38	97	4.7	401	2 AI2255	two-component sens
39	97	4.7	716	2 T02311	hypothetical prote
40	97	4.7	1301	2 S51323	SAC3 protein - yea
41	96.5	4.7	787	2 H90543	conserved hypothet
42	96.5	4.7	897	2 G02529	dynein heavy chain
43	96.5	4.7	4644	1 A38905	probable membrane
44	96	4.6	383	2 S6729	hypothetical prote
45	96	4.6	759	2 T42652	hypothetical prote

ALIGNMENTS

RESULT 1

I75615  
mammary tumor integration site 6 oncogene protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999  
C:Accession: I75615; I75614  
R:Marchetti, A.; Buttitta, F.; Miyazaki, S.; Gallahan, D.; Smith, G.H.; Callahan, R.  
J. Virol. 69, 1932-1938, 1995  
A:Title: Int-6, a highly conserved, widely expressed gene, is mutated by mouse mammary  
A:Reference number: I56914; MUID:95156630; PMID:7853537  
A:Accession: I75615  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-396 <MAR1>  
A:Cross-references: GB:S75224; NID:g913340; PID:g913341  
A:Accession: I75614  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 235-268, 'LKIKFFQSKGIRYVNOAVSGFMEIHRAAVGRPEGHRLGDPLEGNG' <MAR2>  
A:Cross-references: GB:S75223; NID:g913338; PIDN:AAC00047.1; PID:g913339  
A:Note: hypothetical mutant fusion protein  
C:Genetics:  
A:Gene: Int-6

Query Match	100.0%;	Score 2071;	DB 2;	Length 396;
Best Local Similarity	100.0%;	Pred. No. 4.9e-147;		
Matches 396;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MVDFAMDVYKNLYSDDIPHALREKRTTVAQIKQLQAETEPVVKMFEDPETTRQMQSTRD	60	
Db	1	MVDFAMDVYKNLYSDDIPHALREKRTTVAQIKQLQAETEPVVKMFEDPETTRQMQSTRD	60	
Qy	61	GRMLFDYLDKDHGPRQEYLDLTLYRYAKFQYECGNYSGAAEYLYFFRVLPVATDRNALSSL	120	
Db	61	GRMLFDYLDKDHGPRQEYLDLTLYRYAKFQYECGNYSGAAEYLYFFRVLPVATDRNALSSL	120	
Qy	121	WGKLASEILMQNDAAMEDLTRKETIDNNSSVSPQLSQLOQRTWLIIHWSLFFVFNHPKGR	180	
Db	121	WGKLASEILMQNDAAMEDLTRKETIDNNSSVSPQLSQLOQRTWLIIHWSLFFVFNHPKGR	180	
Qy	181	DNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKQVRKRQVLKDLVKVIOQESYTYK	240	
Db	181	DNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKQVRKRQVLKDLVKVIOQESYTYK	240	
Qy	241	DPITEFVECLVYNFDFDGAQKLCRECESVLNDFFLVACLDEFIENARLFIFETFCRIHQ	300	
Db	241	DPITEFVECLVYNFDFDGAQKLCRECESVLNDFFLVACLDEFIENARLFIFETFCRIHQ	300	
Qy	301	CISINMLADKLNMTPEEAERWIVNLIRNARLDADAKIDSKLGHVWGNNAVSPVQOVIETKTK	360	
Db	301	CISINMLADKLNMTPEEAERWIVNLIRNARLDADAKIDSKLGHVWGNNAVSPVQOVIETKTK	360	





Db 179 -SATWCKFASIEILTVDMGDGMEELGKLRWMDVSKSFQDSAVQLRNRTWLLHWSLFLFNH 237  
QY 177 PKGRDNIIDLFLYQOYLNAIQWCMCHILRYLTAVITNKD-----VRKRRQV-----LKD 227  
Db 238 ANGCDTLCDLFFFTY-PLNTIQTSCFWLLRYLTAVVVTNQNNANQKPRNPROSQYQRMRD 296  
QY 228 LKVKIQOESYTKDPITEFVECLVYVNFDPDGAQKLRCECVSLVNDFFLVACLEDFIENA 287  
Db 297 LVRILISQENVEYSDPTSFISALYTEVDPEKAQKLRCECEVLKTDFFLVSLCDHFLEGA 356  
QY 288 RLPIFTFCRIHOCISINMLADKLNMTPEAEARWIVN----- 324  
Db 357 RKLAEAYCRIHVSIVSDVLANKLEMDSAQLIQLVENRNPNPSVAAASNAVAADQSTDESI 416  
QY 325 -----LIRNAR-----LDKIDSKLGHVVMGNNAVSPYQOVIK 358  
Db 417 ESTSTNVVADDLITEATEAEPEPEVQGFQKAKLDGE--SIIIEHPTYSAFQOIIDR 474  
QY 359 TKSLSPRSQWLAMNIEKKLNQ 379  
Db 475 TKSLSPESONLEQSLAKSISE 495

## RESULT 5

T50488  
translation initiation factor eIF-6 [imported] - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jul-2004  
A:Accession: T50488  
R:Crane, R.F.; Craig, R.B.; Norbury, C.J.  
submitted to the EMBL Data Library, January 1999  
A:Description: A conserved fission yeast homolog of the mammalian oncoprotein and eIF3  
A:Reference number: Z25090  
A:Accession: T50488  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-501 <CRA>  
A:Cross-references: UNIPROT:O94513; EMBL:AF117648; PIDN:AAD29969.1  
A:Experimental source: strain 972h(-)  
C:Genetics:  
A:Gene: int6

Query Match 36.7%; Score 760; DB 2; Length 501;

Best Local Similarity 36.7%; Pred. No. 4.8e-49;

Matches 162; Conservative 79; Mismatches 132; Indels 68; Gaps 9;

QY 1 MYDFAMDVYKNLY-----SDIPHALREKRTTVVAQLKQLOAETPIVKMFPEDPETTRMQ 56  
Db 61 MTDYVANLNTNLHGCHTDEDMANAFKRSVLQELSEEEVQGLGVLENPDILTAALR 120  
QY 57 STRDGRMLFDYLDKRGFRQEVLDTLRYAKFOYECGNYSGAAYLYFFRVLPATDRNA 116  
Db 121 --QDKQNGLOHLQEHYNTPIERIAVLYKFAQFYNGCYGGASDLLYHFRFASKDPFLNA 178  
QY 117 LSLSLACKLASEILMQWMDAMELTLKETIDNNSSVSPLOSLQOQRTWLHWSLFFVFNH 176  
Db 179 -SATWCKFASIEILTVDMGDGMEELGKLRWMDVSKSFQDSAVQLRNRTWLLHWSLFLFNH 237  
QY 177 PKGRDNIIDLFLYQOYLNAIQWCMCHILRYLTAVITNKD-----VRKRRQV-----LKD 227  
Db 238 ANGCDTLCDLFFFTY-PLNTIQTSCFWLLRYLTAVVVTNQNNANQKPRNPROSQYQRMRD 296  
QY 228 LKVKIQOESYTKDPITEFVECLVYVNFDPDGAQKLRCECVSLVNDFFLVACLEDFIENA 287  
Db 297 LVRILISQENVEYSDHTVSFNALYTEVDPEKAQKLRCECEVLKTDFFLVSLCDHFLEGA 356  
QY 288 RLPIFTFCRIHOCISINMLADKLNMTPEAEARWIVN----- 324  
Db 357 RKLAEAYCRIHVSIVSDVLANKLEMDSAQLIQLVENRNPNPSVAAASNAVAADQSTDESI 416  
QY 325 -----LIRNAR-----LDKIDSKLGHVVMGNNAVSPYQOVIK 358

## RESULT 7

Db 417 ESTSTNVVADDLITEATEAEPEPEVQGFQKAKLDGE--SIIIEHPTYSAFQOIIDR 474  
QY 359 TKSLSPRSQWLAMNIEKKLNQ 379  
Db 475 TKSLSPESONLEQSLAKSISE 495

## RESULT 6

S56816  
GTPase-activating protein GYP6 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein J1202; protein YJL044C  
C:Species: Saccharomyces cerevisiae  
C:Date: 08-Jul-1995 #sequence\_revision 08-Sep-1995 #text\_change 09-Jul-2004  
A:Accession: S56816; S30061  
R:Fohr, T.M.; Aljinovic, G.  
submitted to the Protein Sequence Database, September 1995  
A:Reference number: S56793  
A:Accession: S56816  
A:Molecule type: DNA  
A:Residues: 1-458 <TOV>  
A:Cross-references: UNIPROT:P32806; EMBL:Z49319; NID:gl008172; PIDN:CAA89335.1; PID:gl0  
R:Strom, M.; Vollmer, P.; Tan, T.J.; Gallwitz, D.  
Nature 361, 736-739, 1993  
A:Title: A yeast GTPase-activating protein that interacts specifically with a member of  
A:Reference number: S30061; MUID:93180913; PMID:8441469  
A:Accession: S30061  
A:Molecule type: DNA  
A:Residues: 1-412, 'L', 414-458 <STR>  
A:Cross-references: EMBL:X68506; NID:g441264; PIDN:CAA48518.1; PID:g5521  
C:Genetics:  
A:Gene: SGD:GYP6  
A:Cross-references: SGD:S0003580; MIPS:YJL044C  
A:Map position: 10L  
C:Superfamily: Saccharomyces cerevisiae GTPase-activating protein GYP6

Query Match 5.6%; Score 116; DB 2; Length 458;

Best Local Similarity 19.1%; Pred. No. 0.49;

Matches 82; Conservative 64; Mismatches 114; Indels 170; Gaps 21;

QY 14 SDDIPHALREKRTTVVAQLKQLOAETPIVKMFPEDPETTRMQMSTDRGMLFY----- 67  
Db 135 SDEKPLTLRETLEIIDLDSRLMD-----DIFOEPKVAQMR-----QLLYNYLLIHQS 184  
QY 68 --LADKRGFRQEVLDTLRYAKFOYECGNYSGAAYLYFFRVLPATDRNALSSLGKLA 135  
Db 185 EHLQYKQGF-HBILSVIYLQLYHGTDLN-----TDLQNVLIIIFNKL 226  
QY 126 SEI-----LMQWMD-----AAMEDL-----TRLKETIDNNSSVSP 156  
Db 227 NQIEPIFYNEENLINWKRVPYTKIFRICLPDLFSKYVYQPPKGTGSKKKKNDHLIHSNLI 286  
QY 157 QSLQORTWLHWSLFFV-----FNHPKGRDNIIDLFLYQOYLNAIQ 198  
Db 287 -----WLIRTRLLFLRELPLKYVLIYVDHVLITENY-----LDIFI----- 323  
QY 199 TWCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTKDPITEFVECLVYVNFDPDG 258  
Db 324 -ACTIITLLS-----IYDELHVLVSQGYEHTNNNDFFEVE-LILHF----- 363  
QY 259 AOKKLRCECVSLVNDFFLVACLEDFIENARLFIETFCRIHOCISINMLADKLNMTPEEA 318  
Db 364 --KKIFEKEDASKDD-----EKFLD-----LCKV-----TGNLC 390  
QY 319 ERWI-----VNLIRNARLDKIDSKLGHVVMGNNA-----VSPYQOVIKTSLSFRSOML 369  
Db 391 ELWYGNKYDDMLICDTPFINAKFGIKTSVLSNETAKLTIDPNRQSL--NKLRRVRQT 448  
QY 370 AMNIEKKLNQ 379  
Db 449 ILKNKKKISQ 458



QY 205 LRYLTAVITNKKVRRQVLKDLVKVIQQESYTKDP-----ITEFVECLYNFDGDA 259  
||| : : : ||| : : | : | : | : | : |  
Db 216 LRHLSVFSSIEEIFKSCNNLKVI-----FTLANPAHEAIKPLEDVKFNLFSSSE 267  
||| : : : ||| : : | : | : | : | : |  
QY 260 QKLRECESVLNVDFLVLACLEDFIENA-----RLFIETFCRIHQICSI 304  
||| : : : ||| : : | : | : | : | : |  
Db 268 RLKI-----YAVADVALAKSTNTLETAASTPMIVAYKVNLISFFIRILIKIKYVTLI 322  
||| : : : ||| : : | : | : | : | : |  
QY 305 NMLADKLNMTPEEAE-RWIVNLIRNARLDKIDS KLGHVVMGNNAVSYPQOVIETK--- 360  
||| : : : ||| : : | : | : | : | : |  
Db 323 NIADK-EIIPEFIQNCRANLISN-----KLQELLF--NSKKAYEQVIESOKILO 370  
||| : : : ||| : : | : | : | : | : |  
QY 361 SLSPRSQWLAMNIE 374  
||| : : : ||| : : | : | : | : | : |  
Db 371 QLGFKNRLQLNSE 384  
||| : : : ||| : : | : | : | : | : |

RESULT 10  
G97089  
Gamma-glutamylcysteine synthetase [imported] - Clostridium acetobutylicum  
C;Species: Clostridium acetobutylicum  
C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: G97089  
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: G97089  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-481 <CUR>  
A;Cross-references: UNIPROT:Q97IV1; GB:A8001437; PIDN:AAK79506.1; PID:g15024489; GSPDB:  
A;Experimental source: Clostridium acetobutylicum ATCC824  
C;Genetics:  
A;Gene: CAC1539  
C;Superfamily: glutamate-cysteine ligase

Query Match 5.2%; Score 108; DB 2; Length 481;  
Best Local Similarity 19.5%; Pred. No. 2.1;  
Matches 88; Conservative 72; Mismatches 136; Indels 156; Gaps 22;

QY 5 AMDVYK--NLYSDDIPHALREKRTTVVAQLKLOAETEPIVVKMFEDPETTRQMSTRDGR 62  
||| : : : ||| : : | : | : | : | : |  
Db 78 AEEYKALNDINNEVNKALKGELLWPLSMPPRLPKEEDIPVAQFPDTEGROKQIVRNG- 136  
||| : : : ||| : : | : | : | : | : |  
QY 63 MLFDYLADKHGFPROEYLDLTYRAKFQYECNYSGAELYLYFF--RVLPATDRNALSSL 120  
||| : : : ||| : : | : | : | : | : |  
Db 137 -----LAIRYGKKQWMISGIHY-----NFSFKDMIDFYROLRIEKTKRKQPIDEM 182  
||| : : : ||| : : | : | : | : | : |  
QY 121 WGLASBILMQNWDAAMEDTLRLKETIDNNSSVSLQSQORTWLHIWSLVFEPFNHPKGR 180  
||| : : : ||| : : | : | : | : | : |  
Db 183 YPSLTRNFLR-----YHWILIYFGASPIC 207  
||| : : : ||| : : | : | : | : | : |  
QY 181 DNIIIDLFLYQPYQLNAIQTMCPHIL-----RYLTAVI-----TNKVRK-----R 221  
||| : : : ||| : : | : | : | : | : |  
Db 208 DSTVNSVIPKE--LEKIEKCPHCAGIKNFNYATSLSRVSRFGSYDTDEKYYTVYFNSL 265  
||| : : : ||| : : | : | : | : | : |  
QY 222 RQVLKDLVKVIQQES--YT----YKDPI-----TEFVECLYNFDFDGAQKKL 263  
||| : : : ||| : : | : | : | : | : |  
Db 266 REYETKIKGMETESNKYSKLGITYKGQVQLNGNLLQSESEF-----YAPIRFKRNIKKG 321  
||| : : : ||| : : | : | : | : | : |  
QY 264 RECESVLNV--DPFLVACLE-----DFIENARLFIETFCRIHQICISNM 306  
||| : : : ||| : : | : | : | : | : |  
Db 322 ETQITALNVNGVYEIRILDVNPFPKVGISVEQMNPLQ-----VFNVFLFEESKSID- 375  
||| : : : ||| : : | : | : | : | : |  
QY 307 LADKANMTPEEAERWIN-----LIRNARL-----DAKIDSK-LGHVVMGNNAVSPY 352  
||| : : : ||| : : | : | : | : | : |  
Db 376 -----EQMERINTNHQAALLGNEDLMLYKYNDSDSRIPLKNFG----- 415  
||| : : : ||| : : | : | : | : | : |  
QY 353 QQVIEKTSLSPRSQWLAMNIEKLNNGRSE 384  
||| : : : ||| : : | : | : | : | : |

A:Molecule type: DNA  
A:Residues: 1-1146 <STO>  
A:Cross-references: UNIPROT:Q910B0; GB:AE004701; GB:AE004806; PIDN:AG0612  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA2732

Query Match 5.1%; Score 106; DB 2; Length 1146;  
Best Local Similarity 23.6%; Pred. No. 9.2; Mismatches 54; Gaps 15;  
Matches 65; Conservative 50; Indels 106; Indels 54; Gaps 15;

QY 17 IPHALREKRTTVA---QLKQLOAETEPVIMFED--PETTRQMOSTRDGRML--FDYL 68  
DB 323 IQHSAGSGKNTIAMLAHQVELRSVADPLMPQFDSIIITDRLALDIQARTIKGYDHV 382

QY 69 ADKHGF---RQYLDTLVR-----YAKFOY---ECNYSGAABLYFFRVLVPATDR 114  
DB 383 ASIFGHSNQAQLRDLVRGRKIIIVTVQKFFPFIIEIGDLSGRT-----FALLID---- 433

QY 115 NALSSLWGLASEILMQNWDAMEDLTRKETIDNNSVSSPLQS-LQOETWLIHWSLFPV 173  
DB 434 EAHSSQGGKTTAR--MHEALSGKTDEEAFEDATQAVNTEIEKRIQSRMLSNASYFAP 491

QY 174 FNHFKGRDNIIDFLYQPOLYNAIOTMCPHILRYLTAVITNKDVKRRQVLKD--LVKV 231  
DB 492 TATPKNR--TSLFGEKTVDDKVEPSPEELTYTT-----KQAIQEGFILDV 537

QY 232 IQQESYTYKDPITEFVECLYVNFDFGAQ--KKLR 264  
DB 538 IAH--YTTLDSFVQAKTVENDPDPDKARALKKIR 570

RESULT 13  
AD1672  
Smc protein essential for chromosome condensation and partition homolog smc [imported]  
C:Species: *Listeria innocua*  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AD1672  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Trieretz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative Genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AD1672  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1186 <GLA>  
A:Cross-references: UNIPROT:Q92AK4; GB:AL592022; PIDN:CAC97148.1; PID:G16414419; GSPDB:Q  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: smc

Query Match 5.1%; Score 105.5; DB 2; Length 1186;  
Best Local Similarity 20.8%; Pred. No. 10;  
Matches 90; Conservative 65; Mismatches 177; Indels 107; Gaps 18;

QY 6 MDVYKNLYSDDIPHALREKRTTVVAQLKQLOAETEPVIMFEDPETTRQMOSTRDGRMLF 65  
DB 452 MEIY-----REVQOTLAKQEAIVGTERALYKH---ETVQMKSRKE---TL 493

QY 66 DYLDKXH-GFRQYBLDTLYRYAKFQYECNYSGAABLYFFRVLVPATDRNALSSLWGL 124  
DB 494 EELADDYAGFFQGVREVL---KAKKEIPGIHGAELVEL---IDIPAKYQAMETALGAS 545

QY 125 ASILMQNWDAMEDLTRKETIDNNSVSSPLQSLOQETWLIHWSLFPVFNHFKGRDNI 184  
DB 546 AQNVVVEDDRVAREATSFLLKTKSGRATPLPLTIQPR-----ELPAATKNAL 593

QY 185 DLFLYQPOY-----LNAIQTWCPIHRLYLTAVITNKDVKRRQVLKDVLV----KVIO 233

DB 594 S---NQPAFIALASEVISFDQKQVSPVILNALTGTTILA-KDL-KGANTLARLVNFRYRIVT 648  
QY 234 QES-----YTKDPITEFVECL-----YVNFDFDGAQ 260

DB 649 LEGDVNAGSGMTGGATKGGKSSILTRKHELQOLAEKTAELNNSSTRELESAYQVAKDSMS 708

QY 261 KKLRECESVLV-----NDPFLVACLEDFTENARLFFTFETCRHQICISINML--ADKL 311  
DB 709 KKRELETRVIGENLRLOKEKELCKLDRETN-----LERENKQLQLYDIEKADGSEEL 763

QY 312 NMTPEAEERWVNLNRNARLDAKIDSKLGHVVMGNNAVSPYQOVIEKTKSLSPRSOMLAW 371  
DB 764 NKLLERKETLQEQVEIAKQIEKTDDEETKAMTSSKALES-----KRTADLSLSLKAQ 818

QY 372 NIEKLNQNSRSEAPNWT 390  
DB 819 IAKREQLQSAVEADVRT 837

RESULT 14  
S64123  
Hypothetical protein YGL113w - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: hypothetical protein G2980  
C:Species: *Saccharomyces cerevisiae*  
C>Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 09-Jul-2004  
C:Accession: S64123; S64121; S69417  
R:Lauguin, G.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64122  
A:Accession: S64123  
A:Molecule type: DNA  
A:Residues: 1-668 <LAU>  
A:Cross-references: UNIPROT:P53135; EMBL:Z72635; NID:G1322662; PIDN:CAA96821.1; PID:G13  
R:Castagnoli, L.; Paoluzi, S.; Minenkova, O.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64112  
A:Accession: S64121  
A:Molecule type: DNA  
A:Residues: 492-668 <CAS>  
A:Cross-references: EMBL:Z72635; MIPS:YGL113w  
A:Experimental source: strain S288C  
R:Castagnoli, L.; Paoluzi, S.; Minenkova, O.  
submitted to the EMBL Data Library, April 1996  
A:Reference number: S69417  
A:Accession: S69417  
A:Molecule type: DNA  
A:Residues: 492-668 <CAW>  
A:Cross-references: EMBL:X97644; NID:G1310710; PIDN:CAA66239.1; PID:G1310711  
C:Genetics:  
A:Gene: SGD:SLD3  
A:Cross-references: SGD:S0003081  
A:Map position: 7L  
C:Superfamily: *Saccharomyces cerevisiae* hypothetical protein YGL113w

Query Match 5.1%; Score 105; DB 2; Length 668;  
Best Local Similarity 20.8%; Pred. No. 5.4;  
Matches 91; Conservative 72; Mismatches 144; Indels 130; Gaps 21;

QY 9 YKNLYSDDIPHALREKRTTVVAQLKQLOAETEPVIMFEDPETTRQMOSTRDGRMLF 56  
DB 77 YFNGYS-----LPERTEVTVTVERAKRETLLKTWDELKFKELLHLWSEEPKSGCKLE 130

QY 57 STRDGRM-----LFDYLDKXHGFRQYBLDTLYRYAKFQYECNYSGAABLYF 104  
DB 131 KDKDLKMLNPPDMKSGKINDYSDP-----KEYIESKYDALFSIH-----TPLAYF 179

QY 105 FR---VLVPATDRNALSSLWGLASEILMQNWDAMEDLTRKETIDNNSV-----SSPLQ 157  
DB 180 VKSNLVRKNTCKTKYGSYSKIAYQAMLOKF---LLSIVQFKDRHNDRLLEPFSFPIA 236

QY 158 SLOQRTWLIHWSLFPVFNHFKGRDNIIDFLYQPOLYNAIOTMCPHILRYLTAVITNKD 217

```
Db 237 DEKRKNCL---TKFVIQDENKNSSTIADLCVWLKSREIKLQ-----ILLLEIIGLNDLD 288
QY 218 -----VRKRQVL--KDLVKVYIQESYTYKDPITBEFVECLYVNFDFGAKX 262
Db 289 WNFPRDEKKYKLLKKRSLMLTKGLVRRRSKKTKTSEK---KGIERITTSLDY----- 339
QY 263 LRECESVLNDFEL-VACL-----EDFIENARLFIFETFCRIHQICISINMLADK 310
Db 340 ---CEQL---DLYLDRACILDLILLSSETPNPDAIEASNGTIQE-----HK 378
QY 311 LNMTPBEAEKWIYNLARNLDAKIDSKLGHVYVGNNAVSPYQOVIEKTKSLSFRSQMLA 370
Db 379 KNILDKSKEASLVGFI-NYVLIFPFYFNKKVPHAV-----EFIQKLKGPSMRPKRAL 428
QY 371 MNIEKLNQNSRSEAPN 387
Db 429 -----KKVNDSTNVSSPN 441
```

## RESULT 15

```
JC7595
scavenger receptor with C-type lectin type I - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C:Accession: JC7595
R:Nakamura, K.; Funakoshi, H.; Miyamoto, K.; Tokunaga, F.; Nakamura, T.
Biochem. Biophys. Res. Commun. 280, 1028-1035, 2001
A:Title: Molecular cloning and functional characterization of a human scavenger receptor
A:Reference number: JC7595; MUID:21092718; PMID:11162630
A:Contents: Placenta
A:Accession: JC7595
A:Molecule type: mRNA
A:Residues: 1-742 <NAK>
A:Cross-references: UNIPROT:Q9BYH7; DDBJ:AB038518
C:Comment: This receptor, a member of the scavenger receptor family, belonging to the tyrosine kinase family, is an important role in host defense. It forms a timer and plays a role in recognizing infected cells.
C:Genetics:
A:Gene: src1-I
A:Map position: 18p11.32
C:Keywords: coiled coil; glycoprotein; transmembrane protein
F:1-39/Domain: cytosolic (amino-terminus) #status predicted <CYT>
F:16-19/Region: internalization signal YKRF
F:40-56/Domain: transmembrane #status predicted <TMM>
F:57-112/Domain: extracellular #status predicted <EXT>
F:113-335/Domain: coiled coil #status predicted <COC>
F:369-384/Region: serine/threonine-rich #status predicted
F:443-589/Domain: collagen-like #status predicted <COL>
F:607-732/Domain: C-type lectin/carbohydrate recognition domain #status predicted <CRD>
```

```
Query Match 5.0%; Score 103.5; DB 2; Length 742;
Best Local Similarity 21.2%; Pred. No. 8;
Matches 76; Conservative 52; Mismatches 115; Indels 115; Gaps 14;
QY 6 MDVYKNLYSDDIPHALREKRTTVVAOLKQLOAETEPVVKMFDEPTTQMQSTRDGRMLF 65
Db 72 METSRQTYDD-----KLTAVESDLKLGDT-----GKKAI 102
QY 66 DYLAADKHGFQRYEYLDLYRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSLGKLA 125
Db 103 STNSELSTFRSDILDLRQ-----LREITEKTSKN-----KDT 135
QY 126 SEILMQNWDAMEDLTRLKETIDNNS-----VSSPIQ-----SLQORTWLIHWSL-FV 172
Db 136 LEKLGASGDALVDROSQLKTELENSFLITTVNKTQAYNGYVTVNLQDQTSVLQGNLQ 195
QY 173 FENHPKGRDNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKDYKRRQVLDLVKVI 232
Db 196 MYSHNVVIMNININLTQVQQRN-----LITN-----LQRSVDDTSQAI 234
QY 233 QSESITYTKDPITBEFVECLYVNFDFGAKKRECESVLVNDFELVACLEDIE--NARLF 290
Db 235 QRIKNDQNLQVVF---LQAKKDTWLKEKVQSLQTLAANNSALAKANNNTLEDNNSQLN 291
```

```
QY 291 IF-----ETFCRIHQICISINM-----LADKLNMTPEAEERW---IVNLIEN 328
Db 292 SFTGQMNITTTISQANEQNKLQDLHLKDAENRTAKFNQLEERFQLFETDIVNIISN 349
```

Search completed: March 8, 2005, 20:08:09  
Job time : 50 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 20:00:57 ; Search time 43 Seconds  
(without alignments)  
687.466 Million cell updates/sec

Title: US-10-783-415-4  
Perfect score: 2071  
Sequence: 1 MVDFAWDYKLYNDDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2071	100.0	396	3	US-08-875-847B-2
2	2071	100.0	396	3	US-08-875-847B-4
3	2071	100.0	396	3	US-09-378-842-2
4	2071	100.0	396	3	US-09-378-842-4
5	2071	100.0	396	4	US-09-858-152B-2
6	2071	100.0	396	4	US-09-858-152B-4
7	2071	100.0	445	4	US-09-976-594-841
8	106	5.1	1575	4	US-09-917-254-83
9	106	5.1	1575	4	US-09-949-016-6743
10	104	5.0	489	4	US-09-248-796A-19030
11	101.5	4.9	313	4	US-09-248-796A-23197
12	101.5	4.9	578	1	US-08-766-014-4
13	101.5	4.9	608	1	US-08-766-014-3
14	100.5	4.9	1196	4	US-09-107-532A-3944
15	100	4.8	1559	4	US-09-949-016-10190
16	98.5	4.8	872	1	US-08-766-014-2
17	98.5	4.8	1056	4	US-09-079-030-217
18	98.5	4.8	2663	4	US-09-538-092-1252
19	98.5	4.8	4536	4	US-09-180-422B-27
20	98.5	4.8	4536	4	US-09-079-030-1
21	98.5	4.8	4563	4	US-09-538-092-842
22	98	4.7	2662	4	US-09-595-684B-31
23	97.5	4.7	1979	4	US-09-949-016-6468
24	97.5	4.7	2047	4	US-09-949-016-7404
25	97.5	4.7	4563	4	US-09-108-006C-1
26	96.5	4.7	2807	4	US-09-543-681A-4980
27	95	4.6	1011	4	US-09-489-039A-11808

28	94.5	4.6	789	4	US-09-248-796A-19294	Sequence 19294, A
29	94.5	4.6	913	3	US-08-971-089-4	Sequence 4, Appli
30	94	4.5	738	4	US-09-107-532A-6218	Sequence 6218, Ap
31	93.5	4.5	721	4	US-08-851-435-2	Sequence 2, Appli
32	93.5	4.5	746	4	US-08-851-435-6	Sequence 6, Appli
33	93	4.5	420	4	US-09-270-767-42980	Sequence 42980, A
34	93	4.5	599	2	US-08-910-551B-2	Sequence 2, Appli
35	93	4.5	916	4	US-09-688-078-9	Sequence 9, Appli
36	92.5	4.5	390	4	US-09-107-532A-4359	Sequence 4359, Ap
37	92.5	4.5	530	4	US-09-270-767-41545	Sequence 41545, A
38	92.5	4.5	792	4	US-09-543-681A-6617	Sequence 6617, Ap
39	92	4.4	660	4	US-09-248-796A-14578	Sequence 14578, A
40	92	4.4	991	4	US-09-248-796A-15239	Sequence 15239, A
41	91.5	4.4	572	4	US-09-134-000C-4237	Sequence 4237, Ap
42	90.5	4.4	568	3	US-09-134-001C-3768	Sequence 3768, Ap
43	90.5	4.4	1753	4	US-09-248-796A-19154	Sequence 19154, A
44	90	4.3	655	4	US-09-949-016-6759	Sequence 6759, Ap
45	90	4.3	669	4	US-09-949-016-10241	Sequence 10241, A

ALIGNMENTS

RESULT 1  
US-08-875-847B-2  
; Sequence 2, Application US/08875847B  
; Patent No. 6255105  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United  
; APPLICANT: States of America as represented by the  
; APPLICANT: Secretary, Department of Health and Human  
; APPLICANT: Services, Callahan, Robert; Marchetti,  
; APPLICANT: Antonio; Butticita, Fiamma; Smith, Gilbert H.  
; TITLE OF INVENTION: Nucleotide And Deduced  
; TITLE OF INVENTION: Amino Acid Sequences Of A New Tumor Gene,  
; TITLE OF INVENTION: Int6, And The Use Of Reagents Derived From  
; TITLE OF INVENTION: These Sequences In Diagnostic Assays.  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: MS WORD 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/875,847B  
; FILING DATE: 09-FEB-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/385,998  
; FILING DATE: 09-FEB-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: William S. Feller  
; REGISTRATION NUMBER: 26,728  
; REFERENCE/DOCKET NUMBER: 2026-4179PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 396 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown





OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MS WORD 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/378,842  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/875,847  
FILING DATE: 09-FEB-1996  
APPLICATION NUMBER: 08/385,998  
FILING DATE: 09-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: William S. Feiler  
REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2026-4179PCT  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 396 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
US-09-378-842-2

Query Match 100.0%; Score 2071; DB 3; Length 396;  
Best Local Similarity 100.0%; Pred. No. 2.8e-212;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVDFAVDYKKNLYSDDI PHALREKRTTVAQKQLQAEITEPIVKMFEDPETTRQMOSTRD 60  
DB 1 MVDFAVDYKKNLYSDDI PHALREKRTTVAQKQLQAEITEPIVKMFEDPETTRQMOSTRD 60  
QY 61 GRMLFDYLDKXGFRQEYLDITLYRYAKFOYECGNYSGAAEYLYFFRVLVLPATDRNALSSL 120  
DB 61 GRMLFDYLDKXGFRQEYLDITLYRYAKFOYECGNYSGAAEYLYFFRVLVLPATDRNALSSL 120  
QY 121 WGLKASEILMQNWDAAMEDLTRKKTIDNNSSVSPQLSQRTWLIIHWSLFFVFNHPKGR 180  
DB 121 WGLKASEILMQNWDAAMEDLTRKKTIDNNSSVSPQLSQRTWLIIHWSLFFVFNHPKGR 180  
QY 181 DNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQQSSYTYK 240  
DB 181 DNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQQSSYTYK 240  
QY 241 DPITEFVECLYVDFDGAQKKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300  
DB 241 DPITEFVECLYVDFDGAQKKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300  
QY 301 CISINMLADKLNMTPEEAERWIVNIRNARLDAKIDSKLGHVVMGNNVSPYQQVIEKTK 360  
DB 301 CISINMLADKLNMTPEEAERWIVNIRNARLDAKIDSKLGHVVMGNNVSPYQQVIEKTK 360  
QY 361 SLFSRQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396  
DB 361 SLFSRQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 4  
US-09-378-842-4  
Sequence 4, Application US/09378842  
Patent No. 6342392  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America as represented by the  
APPLICANT: Secretary, Department of Health and Human Services; Callahan, Robert; Marchetti,  
APPLICANT: Antonio; Buttitta, Fiamma; Smith, Gilbert H.  
TITLE OF INVENTION: Nucleotide And Deduced  
TITLE OF INVENTION: Amino Acid Sequences Of A New Tumor Gene,  
Int6, And The Use Of Reagents Derived From

TITLE OF INVENTION: These Sequences In Diagnostic Assays,  
TITLE OF INVENTION: Vaccines, Immunotherapy And Gene Therapy  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MS WORD 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/378,842  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/875,847  
FILING DATE: 09-FEB-1996  
APPLICATION NUMBER: 08/385,998  
FILING DATE: 09-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: William S. Feiler  
REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2026-4179PCT  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 396 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
US-09-378-842-4

Query Match 100.0%; Score 2071; DB 3; Length 396;  
Best Local Similarity 100.0%; Pred. No. 2.8e-212;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVDFAVDYKKNLYSDDI PHALREKRTTVAQKQLQAEITEPIVKMFEDPETTRQMOSTRD 60  
DB 1 MVDFAVDYKKNLYSDDI PHALREKRTTVAQKQLQAEITEPIVKMFEDPETTRQMOSTRD 60  
QY 61 GRMLFDYLDKXGFRQEYLDITLYRYAKFOYECGNYSGAAEYLYFFRVLVLPATDRNALSSL 120  
DB 61 GRMLFDYLDKXGFRQEYLDITLYRYAKFOYECGNYSGAAEYLYFFRVLVLPATDRNALSSL 120  
QY 121 WGLKASEILMQNWDAAMEDLTRKKTIDNNSSVSPQLSQRTWLIIHWSLFFVFNHPKGR 180  
DB 121 WGLKASEILMQNWDAAMEDLTRKKTIDNNSSVSPQLSQRTWLIIHWSLFFVFNHPKGR 180  
QY 181 DNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQQSSYTYK 240  
DB 181 DNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQQSSYTYK 240  
QY 241 DPITEFVECLYVDFDGAQKKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300  
DB 241 DPITEFVECLYVDFDGAQKKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300  
QY 301 CISINMLADKLNMTPEEAERWIVNIRNARLDAKIDSKLGHVVMGNNVSPYQQVIEKTK 360  
DB 301 CISINMLADKLNMTPEEAERWIVNIRNARLDAKIDSKLGHVVMGNNVSPYQQVIEKTK 360  
QY 361 SLFSRQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396  
DB 361 SLFSRQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

```
RESULT 5
US-09-858-152B-2
; Sequence 2, Application US/09858152B
; Patent No. 6737251
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Fianna
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
; CURRENT APPLICATION NUMBER: US/09/858,152B
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 09/858,152
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Murine INT6
US-09-858-152B-2

Query Match 100.0%; Score 2071; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 2.8e-212;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDYKLYSDDI PHALREKRTTVAQLKQLQAETPIVKMFEDPETTRQMQSTRD 60
Db 1 MVDFAVDYKLYSDDI PHALREKRTTVAQLKQLQAETPIVKMFEDPETTRQMQSTRD 60
Qy 61 GRMLFDYLADKHGFRQEVLDLTRYAKFQYECGNYSGAAEYLFFRVLPVLPATDRNALSSL 120
Db 61 GRMLFDYLADKHGFRQEVLDLTRYAKFQYECGNYSGAAEYLFFRVLPVLPATDRNALSSL 120
Qy 121 WGLASEILMQNWDAAEDLTRLKETIDNNSSVSPQLSQLOQTWLIHWSLFFVFNHPKGR 180
Db 121 WGLASEILMQNWDAAEDLTRLKETIDNNSSVSPQLSQLOQTWLIHWSLFFVFNHPKGR 180
Qy 181 DNIIDFLYQPOLYNAIQTMCPHILRYLTAVITNKDVKRRQVLKOLVKVIQOESYTYK 240
Db 181 DNIIDFLYQPOLYNAIQTMCPHILRYLTAVITNKDVKRRQVLKOLVKVIQOESYTYK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFLVALEDFTIENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFLVALEDFTIENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
Qy 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396
Db 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396

RESULT 6
US-09-858-152B-4
; Sequence 4, Application US/09858152B
; Patent No. 6737251
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Fianna
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
```

```
; CURRENT APPLICATION NUMBER: US/09/858,152B
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 09/858,152
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-152B-4

Query Match 100.0%; Score 2071; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 2.8e-212;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDYKLYSDDI PHALREKRTTVAQLKQLQAETPIVKMFEDPETTRQMQSTRD 60
Db 1 MVDFAVDYKLYSDDI PHALREKRTTVAQLKQLQAETPIVKMFEDPETTRQMQSTRD 60
Qy 61 GRMLFDYLADKHGFRQEVLDLTRYAKFQYECGNYSGAAEYLFFRVLPVLPATDRNALSSL 120
Db 61 GRMLFDYLADKHGFRQEVLDLTRYAKFQYECGNYSGAAEYLFFRVLPVLPATDRNALSSL 120
Qy 121 WGLASEILMQNWDAAEDLTRLKETIDNNSSVSPQLSQLOQTWLIHWSLFFVFNHPKGR 180
Db 121 WGLASEILMQNWDAAEDLTRLKETIDNNSSVSPQLSQLOQTWLIHWSLFFVFNHPKGR 180
Qy 181 DNIIDFLYQPOLYNAIQTMCPHILRYLTAVITNKDVKRRQVLKOLVKVIQOESYTYK 240
Db 181 DNIIDFLYQPOLYNAIQTMCPHILRYLTAVITNKDVKRRQVLKOLVKVIQOESYTYK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFLVALEDFTIENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFLVALEDFTIENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
Qy 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396
Db 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396

RESULT 7
US-09-976-594-841
; Sequence 841, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 841
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 1330056CD1
US-09-976-594-841

Query Match 100.0%; Score 2071; DB 4; Length 445;
Best Local Similarity 100.0%; Pred. No. 3.4e-212;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MVDFAVDVYKNLYSDDI PHALREKRTTVVAQLKQLQAEPIVKMFDPETTRQMSQTRD 60  
DB 50 MVDFAVDVYKNLYSDDI PHALREKRTTVVAQLKQLQAEPIVKMFDPETTRQMSQTRD 109  
QY 61 GRMLFDYLDKGFQROEYLDLYRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSSL 120  
DB 110 GRMLFDYLDKGFQROEYLDLYRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSSL 169  
QY 121 WGLKASEILMOWNDAAEDLTTLKETIDNNSVSSPLOSQORTWLHLSLFFVFNHPKGR 180  
DB 170 WGLKASEILMOWNDAAEDLTTLKETIDNNSVSSPLOSQORTWLHLSLFFVFNHPKGR 229  
QY 181 DNIIDFLYQPOLYNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVYQOESYTYK 240  
DB 230 DNIIDFLYQPOLYNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVYQOESYTYK 289  
QY 241 DPITFVECLYVNFDPDGAQKLRCESVLVNDFLVACLEDFIENARLFIFETFCRIHQ 300  
DB 290 DPITFVECLYVNFDPDGAQKLRCESVLVNDFLVACLEDFIENARLFIFETFCRIHQ 349  
QY 301 CISINMLADKLNMTPEEAERWILIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
DB 350 CISINMLADKLNMTPEEAERWILIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 409  
QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396  
DB 410 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 8  
US-09-917-254-83  
; Sequence 83, Application US/09917254  
; Patent No. 6703204  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George  
; APPLICANT: Baak, Jan  
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer  
; FILE REFERENCE: B0801/7224(JRV)  
; CURRENT APPLICATION NUMBER: US/09/917,254  
; CURRENT FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: US 60/222,093  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 83  
; LENGTH: 1575  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-917-254-83

Query Match 5.1%; Score 106; DB 4; Length 1575;  
Best Local Similarity 21.8%; Pred. No. 0.16;  
Matches 86; Conservative 49; Mismatches 133; Indels 126; Gaps 19;  
QY 48 DPETTRQMSQTRDGRMLFDYLDKGFQROEYLDLYRYAKFOYE-----CGNYSG 97  
DB 318 DPENT-----LLALKKPEAQ--LPVYPPFAAAMYQNELFNQLKQNTMNYLA 361  
QY 98 -----AAEYLYFFRVLVPATDRNALSSLGWKLASEIL-MQNWDAAMEDTLRLKETIDNN 150  
DB 362 HEELIAVEMLSAVALNQALNESDNLVSQNLRSIPAIGLNLDKA-----YVERYANT 415  
QY 151 SVSPSPLOSQORTWLHLSLFFVFNHPKGRDNIIDFLYQPO-----YLN-AIQT 199  
DB 416 LLSVKLEVLSSQODNLSWNEI-----QNCIDMVNAQIQEENDRVVAVGYINEAIDE 466  
QY 200 MCPHILRYLTAVITNKDVRKRQVLKDLVKVYQOESYTYKDPITFVECLYVNFDPGA 259  
DB 467 GNP--LRLETLLTLPANIS-----DVPDPAHQHYQDVL-----YHAK 502  
QY 320 RWIVNLIR-NARLDAKIDSKLGHVVMGN-----NAVSP-----YQOVIETKSL-----SF 364  
DB 537 QWTLVVDVNOCLGKSSDILSVLKSSSTNSANDIIPCADKYYDALVKAKELKSERVSS 596  
QY 365 RSQMLAMNIEKKLN--QNSRSEAPNWTQDSGFY 396  
DB 597 DGSMLKLNHLKKYDYNYNTDSKSSSWTPESCIFY 630

RESULT 10  
US-09-248-796A-19030  
; Sequence 19030, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:

QY 320 RWIVNLIR-NARLDAKIDSKLGHVVMGN-----NAVSP-----YQOVIETKSL-----SF 364  
DB 537 QWTLVVDVNOCLGKSSDILSVLKSSSTNSANDIIPCADKYYDALVKAKELKSERVSS 596  
QY 365 RSQMLAMNIEKKLN--QNSRSEAPNWTQDSGFY 396  
DB 597 DGSMLKLNHLKKYDYNYNTDSKSSSWTPESCIFY 630

RESULT 9  
US-09-949-016-6743  
; Sequence 6743, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6743  
; LENGTH: 1575  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-6743

Query Match 5.1%; Score 106; DB 4; Length 1575;  
Best Local Similarity 21.8%; Pred. No. 0.16;  
Matches 86; Conservative 49; Mismatches 133; Indels 126; Gaps 19;  
QY 48 DPETTRQMSQTRDGRMLFDYLDKGFQROEYLDLYRYAKFOYE-----CGNYSG 97  
DB 318 DPENT-----LLALKKPEAQ--LPVYPPFAAAMYQNELFNQLKQNTMNYLA 361  
QY 98 -----AAEYLYFFRVLVPATDRNALSSLGWKLASEIL-MQNWDAAMEDTLRLKETIDNN 150  
DB 362 HEELIAVEMLSAVALNQALNESDNLVSQNLRSIPAIGLNLDKA-----YVERYANT 415  
QY 151 SVSPSPLOSQORTWLHLSLFFVFNHPKGRDNIIDFLYQPO-----YLN-AIQT 199  
DB 416 LLSVKLEVLSSQODNLSWNEI-----QNCIDMVNAQIQEENDRVVAVGYINEAIDE 466  
QY 200 MCPHILRYLTAVITNKDVRKRQVLKDLVKVYQOESYTYKDPITFVECLYVNFDPGA 259  
DB 467 GNP--LRLETLLTLPANIS-----DVPDPAHQHYQDVL-----YHAK 502  
QY 260 QKKLRCESVLVNDFLVACLEDFIENARLFIFETFCRIHQICISINMLADKLNMTPEEA 319  
DB 503 SQKLGDSSEV-----SKVLWDE-----IQQAVDEANVDEDRAK 536  
QY 320 RWIVNLIR-NARLDAKIDSKLGHVVMGN-----NAVSP-----YQOVIETKSL-----SF 364  
DB 537 QWTLVVDVNOCLGKSSDILSVLKSSSTNSANDIIPCADKYYDALVKAKELKSERVSS 596  
QY 365 RSQMLAMNIEKKLN--QNSRSEAPNWTQDSGFY 396  
DB 597 DGSMLKLNHLKKYDYNYNTDSKSSSWTPESCIFY 630

```

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19030
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19030

```

Query Match 5.0%; Score 104; DB 4; Length 489;  
Best Local Similarity 20.5%; Pred. No. 0.04;  
Matches 75; Conservative 68; Mismatches 139; Indels 84; Gaps 16;

[illegible]

```

RESULT 11
US-09-248-796A-23197
; Sequence 23197, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 23197
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-23197

```

Query Match 4.9%; Score 101.5; DB 4; Length 313;

[illegible]

## RESINT 12

US-08-766-014-4  
; Sequence 4, Application US/08766014  
; Patent No. 5744312  
; GENERAL INFORMATION:  
; APPLICANT: Mamone, Joseph A.  
; APPLICANT: Davis, Maria  
; APPLICANT: Sha, Dan  
; TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/766,014  
; FILING DATE: Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; PRIOR APPLICATION DATA: including application  
; PRIOR APPLICATION DATA: described below:  
; APPLICATION NUMBER: 60/008,688  
; FILING DATE: December 15, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 223/104  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 578 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-766-014-4






GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 20:08:59 ; Search time 44 Seconds  
(without alignments)  
671.842 Million cell updates/sec

Title: US-10-783-415-4  
Perfect score: 396  
Sequence: 1 MVDPMVYKLYSDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table:  Gapop 60.0, Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	396	100.0	396	3	US-08-875-847B-2
2	396	100.0	396	3	US-08-875-847B-4
3	396	100.0	396	3	US-08-378-842-2
4	396	100.0	396	3	US-08-378-842-4
5	396	100.0	396	4	US-09-858-152B-2
6	396	100.0	396	4	US-09-858-152B-4
7	396	100.0	445	4	US-09-976-594-841
8	7	1.8	53	4	US-09-270-767-46143
9	7	1.8	55	4	US-09-270-767-35620
10	7	1.8	55	4	US-09-270-767-50837
11	7	1.8	62	4	US-09-621-976-7296
12	7	1.8	70	4	US-09-621-976-4933
13	7	1.8	74	3	US-09-025-151-8
14	7	1.8	74	4	US-09-637-240-8
15	7	1.8	81	4	US-09-621-976-7295
16	7	1.8	81	4	US-09-621-976-7298
17	7	1.8	94	1	US-08-167-035-23
18	7	1.8	94	1	US-08-208-887A-23
19	7	1.8	94	2	US-08-539-005-23
20	7	1.8	94	3	US-09-280-598-25
21	7	1.8	120	4	US-09-489-039A-10583
22	7	1.8	133	4	US-09-034-418-6
23	7	1.8	148	4	US-09-489-039A-14211
24	7	1.8	172	4	US-09-538-092-450
25	7	1.8	175	4	US-09-248-796A-19585
26	7	1.8	182	2	US-08-467-603-89
27	7	1.8	182	2	US-08-466-793-89

Sequence 89, Appl  
Sequence 89, Appl  
Sequence 36727, A  
Sequence 51944, A  
Sequence 36, Appl  
Sequence 9350, Ap  
Sequence 16553, A  
Sequence 20877, A  
Sequence 522, App  
Sequence 4400, Ap  
Sequence 16, Appl  
Sequence 7, Appl  
Sequence 374, App  
Sequence 1, Appl  
Sequence 2, Appl  
Sequence 83, Appl  
Sequence 83, Appl

28 7 1.8 182 2 US-08-491-861A-89  
29 7 1.8 182 4 US-09-374-871A-89  
30 7 1.8 196 4 US-09-270-767-36727  
31 7 1.8 196 4 US-09-270-767-51944  
32 7 1.8 207 3 US-09-302-769-36  
33 7 1.8 238 4 US-09-489-039A-9350  
34 7 1.8 238 4 US-09-248-796A-16553  
35 7 1.8 259 4 US-09-248-796A-20877  
36 7 1.8 307 4 US-09-710-279-522  
37 7 1.8 308 3 US-09-134-001C-4400  
38 7 1.8 314 3 US-08-653-648A-16  
39 7 1.8 314 4 US-09-564-418-7  
40 7 1.8 373 4 US-09-149-476-374  
41 7 1.8 375 1 US-08-027-986-1  
42 7 1.8 375 1 US-08-027-986-2  
43 7 1.8 403 4 US-08-887-534A-83  
44 7 1.8 403 4 US-09-527-431-83  
45 7 1.8 403 4 US-09-446-861-83

#### ALIGNMENTS

RESULT 1  
US-08-875-847B-2  
; Sequence 2, Application US/08875847B  
; Patent No. 6255105  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United  
; APPLICANT: States of America as represented by the  
; APPLICANT: Secretary, Department of Health and Human  
; APPLICANT: Services; Callahan, Robert; Marchetti,  
; APPLICANT: Antonio; Nuttitta, Fiamma; Smith, Gilbert H.  
; TITLE OF INVENTION: Nucleotide And Deduced  
; TITLE OF INVENTION: Amino Acid Sequences Of A New Tumor Gene,  
; TITLE OF INVENTION: Int6. And the Use Of Reagents Derived From  
; TITLE OF INVENTION: These Sequences In Diagnostic Assays,  
; TITLE OF INVENTION: Vaccines, Immunotherapy And Gene Therapy  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154

COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: MS WORD 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/875,847B  
; FILING DATE: 09-FEB-1996  
; CLASSIFICATION: 435  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 08/385,998  
; FILING DATE: 09-FEB-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: William S. Feiler  
; REGISTRATION NUMBER: 26,728  
; REFERENCE/DOCKET NUMBER: 2026-4179PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 396 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown

US-08-875-847B-2

```
Query Match      100.0%; Score 396; DB 3; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396: Conservative 0; Mismatches 0; Indels
```

Qy	1	MVD	F	A	M	V	N	K	N	L	Y	S	D	I	P	H	A	L	R	E	K	T	T	V	V	A	Q	L	Q	A	E	T	P	I	V	K	M	F	E	D	P	E	T	T	R	Q	M	O	S	T	R	D	60
Db	1	MVD	F	A	M	V	N	K	N	L	Y	S	D	I	P	H	A	L	R	E	K	T	T	V	V	A	Q	L	Q	A	E	T	P	I	V	K	M	F	E	D	P	E	T	T	R	Q	M	O	S	T	R	D	60
Qy	61	GRM	L	P	D	L	A	K	H	G	F	R	Q	E	Y	L	D	T	L	R	Y	A	K	F	Q	E	C	G	N	S	G	A	E	Y	L	F	F	R	V	L	V	P	A	T	R	N	A	L	S	S	L	120	
Db	61	GRM	L	P	D	L	A	K	H	G	F	R	Q	E	Y	L	D	T	L	R	Y	A	K	F	Q	E	C	G	N	S	G	A	E	Y	L	F	F	R	V	L	V	P	A	T	R	N	A	L	S	S	L	120	
Qy	121	WGL	A	S	E	I	L	M	Q	N	D	A	M	E	D	T	L	R	K	E	T	I	D	N	N	S	V	S	P	L	Q	S	L	Q	R	T	W	L	I	H	W	S	L	F	V	F	N	H	P	K	R	180	
Db	121	WGL	A	S	E	I	L	M	Q	N	D	A	M	E	D	T	L	R	K	E	T	I	D	N	N	S	V	S	P	L	Q	S	L	Q	R	T	W	L	I	H	W	S	L	F	V	F	N	H	P	K	R	180	
Qy	181	DNI	I	D	L	F	L	Y	P	O	Y	L	N	A	I	Q	T	M	C	H	I	L	R	Y	L	T	T	A	V	I	N	K	D	V	R	K	R	Q	V	L	K	D	L	V	K	I	O	E	S	Y	T	K	240
Db	181	DNI	I	D	L	F	L	Y	P	O	Y	L	N	A	I	Q	T	M	C	H	I	L	R	Y	L	T	T	A	V	I	N	K	D	V	R	K	R	Q	V	L	K	D	L	V	K	I	O	E	S	Y	T	K	240
Qy	241	DP	T	E	V	E	C	L	Y	N	F	D	G	A	K	L	R	E	C	E	S	V	L	N	D	F	L	V	A	C	L	E	D	F	T	E	N	A	R	L	I	F	I	F	T	F	C	R	I	H	Q	300	
Db	241	DP	T	E	V	E	C	L	Y	N	F	D	G	A	K	L	R	E	C	E	S	V	L	N	D	F	L	V	A	C	L	E	D	F	T	E	N	A	R	L	I	F	I	F	T	F	C	R	I	H	Q	300	
Qy	301	CIS	I	N	L	M	A	D	K	N	T	P	E	A	E	R	V	I	N	L	R	N	A	L	D	A	K	I	D	S	K	L	G	H	V	M	G	N	N	A	V	S	P	Y	O	O	V	E	K	T	K	360	
Db	301	CIS	I	N	L	M	A	D	K	N	T	P	E	A	E	R	V	I	N	L	R	N	A	L	D	A	K	I	D	S	K	L	G	H	V	M	G	N	N	A	V	S	P	Y	O	O	V	E	K	T	K	360	
Qy	361	S	I	S	F	R	O	M	L	A	M	N	I	E	K	L	N	O	N	S	R	E	A	P	N	W	A	T	O	D	S	G	F	Y	396																		
Db	361	S	I	S	F	R	O	M	L	A	M	N	I	E	K	L	N	O	N	S	R	E	A	P	N	W	A	T	O	D	S	G	F	Y	396																		

RESIN, T 2

RESOUR 2  
US-08-875-847B-4  
; Sequence 4, Application US/08875847B  
; Patent No. 6255105

; GENERAL INFORMATION:  
 ; APPLICANT: The Government of the United  
 ; APPLICANT: States of America as represented by the  
 ; APPLICANT: Secretary, Department of Health and Human  
 ; APPLICANT: Services; Callahan, Robert; Marchetti,  
 ; APPLICANT: Antonio; Buttitta, Fiamma; Smith, Gilbert H.  
 ; TITLE OF INVENTION: Nucleoside And Deduced  
 ; TITLE OF INVENTION: Amino Acid Sequences of A New Tumor Gene,  
 ; TITLE OF INVENTION: Int6, And The Use of Reagents Derived From  
 ; TITLE OF INVENTION: These Sequences In Diagnostic Assays,  
 ; TITLE OF INVENTION: Vaccines, Immunotherapy And Gene Therapy  
 ; NUMBER OF INVENTIONS: 32  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
 ; STREET: 345 PARK AVENUE  
 ; CITY: NEW YORK  
 ; STATE: NEW YORK  
 ; COUNTRY: USA  
 ; ZIP: 10154

```
,
, COMPUTER READABLE FORM:
, MEDIUM TYPE: FLOPPY DISK
, COMPUTER: IBM PC COMPATIBLE
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: MS WORD 97
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/875,847B
, FILING DATE: 09-FEB-1996
, CLASSIFICATION: 435
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 08/395,998
, FILING DATE: 09-FEB-1995
, CLASSIFICATION: 435
, ATTORNEY/AGENT INFORMATION:
, NAME: William S. Feiler
```

REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2036-4179PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 396 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
US-08-875-847B-4

Query Match	100.0.0%;	Score 396;	DB 3;	Length 396;
Best Local Similarity	100.0.0%;	Pred. No. 0;		
Matches 396;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MVDFAMDVYKNLYSDDIPHALREKRTTVVAQLKQLAETEPIVKMPEDPETTRMQSTRD	60	
Db	1	MVDFAMDVYKNLYSDDIPHALREKRTTVVAQLKQLAETEPIVKMPEDPETTRMQSTRD	60	
Qy	61	GRMLFDYLDKDHGRQEYDLTLRYAKFOYECGNYSGAAEYLFFRVFLVPATDRLNALSSL	120	
Db	61	GRMLFDYLDKDHGRQEYDLTLRYAKFOYECGNYSGAAEYLFFRVFLVPATDRLNALSSL	120	
Qy	121	WGKLASEILMONDWAAMEDITRLKETIDNNSSVSPQSLQOORTWLHWSLVFVFNHPKGR	180	
Db	121	WGKLASEILMONDWAAMEDITRLKETIDNNSSVSPQSLQOORTWLHWSLVFVFNHPKGR	180	
Qy	181	DNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTYK	240	
Db	181	DNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTYK	240	
Qy	241	DPITEFVECLYVNFDFDGAOKKLRECESVLVNDFFLVACLEDIEIENARLFIETFCCRHHQ	300	
Db	241	DPITEFVECLYVNFDFDGAOKKLRECESVLVNDFFLVACLEDIEIENARLFIETFCCRHHQ	300	
Qy	301	CISINMLADKLANMTPEEAERWIVNLRNARLDAKIDSKLGHVVGNNNAVSPYQOVIKTK	360	
Db	301	CISINMLADKLANMTPEEAERWIVNLRNARLDAKIDSKLGHVVGNNNAVSPYQOVIKTK	360	
Qy	361	SLSFRSQMLAMNTEKKLQNNRSSEAPNWTQDSGFY	396	
Db	361	SLSFRSQMLAMNTEKKLQNNRSSEAPNWTQDSGFY	396	

### RESULT. T 3

US-09-378-842-2  
; Sequence 2, Application US/09378842  
: Patent No. 6342392

; GENERAL INFORMATION:  
 ; APPLICANT: The Government of the United  
 ; APPLICANT: States of America as represented by the  
 ; APPLICANT: Secretary, Department of Health and Human  
 ; APPLICANT: Services; Callahan, Robert; Marchetti,  
 ; APPLICANT: Antonio; Buttrita, Fiamma; Smith, Gilbert H.  
 ; TITLE OF INVENTION: Nucleoside And Deduced  
 ; TITLE OF INVENTION: Amino Acid Sequences of A New Tumor Gene,  
 ; TITLE OF INVENTION: Intg, And The Use of Reagents Derived From  
 ; TITLE OF INVENTION: These Sequences In Diagnostic Assays,  
 ; TITLE OF INVENTION: Vaccines, Immunotherapy And Gene Therapy  
 ; NUMBER OF SEQUENCES: 32  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
 ; STREET: 345 PARK AVENUE  
 ; CITY: NEW YORK  
 ; STATE: NEW YORK  
 ; COUNTRY: USA  
 ; ZIP: 10154  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: FLOPPY DISK  
 ; COMPUTER: IBM PC COMPATIBLE



OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MS WORD 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/378,842  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/875,847  
FILING DATE: 09-FEB-1996  
APPLICATION NUMBER: 08/385,998  
FILING DATE: 09-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: William S. Feiler  
REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2026-4179PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 396 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
US-09-378-842-2

Query Match 100.0%; Score 396; DB 3; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MVDFAVDYKXLYSDDIPHALREKRTTVAQLKQLQAEETPIVKMFEDPETTRQMQSTRD 60  
Db 1 MVDFAVDYKXLYSDDIPHALREKRTTVAQLKQLQAEETPIVKMFEDPETTRQMQSTRD 60  
Qy 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFVRVLVPATDRLNSSL 120  
Db 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFVRVLVPATDRLNSSL 120  
Qy 121 WGLASILMKNQWDAAMEDLTRKETIDNNSVSSPLOSQLOQTWLIHWSLVFFNHPKGR 180  
Db 121 WGLASILMKNQWDAAMEDLTRKETIDNNSVSSPLOSQLOQTWLIHWSLVFFNHPKGR 180  
Qy 181 DNIIDFLYQPYQLNAIQTMCPHILRYLTAVITNKDVRKRRQVLKOLVKVIOQESYTK 240  
Db 181 DNIIDFLYQPYQLNAIQTMCPHILRYLTAVITNKDVRKRRQVLKOLVKVIOQESYTK 240  
Qy 241 DPITEFVECLYVNFDFDGAQKKLRECESLVNDFFLVACLEDFTENARLPIFETFCRIHQ 300  
Db 241 DPITEFVECLYVNFDFDGAQKKLRECESLVNDFFLVACLEDFTENARLPIFETFCRIHQ 300  
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360  
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360  
Qy 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396  
Db 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396

RESULT 4  
US-09-378-842-4  
Sequence 4, Application US/09378842  
Patent No. 6342392  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America as represented by the  
APPLICANT: Secretary, Department of Health and Human Services; Callahan, Robert; Marchetti,  
APPLICANT: Antonio; Buttitta, Fiamma; Smith, Gilbert H.  
TITLE OF INVENTION: Nucleotide And Deduced  
TITLE OF INVENTION: Amino Acid Sequences Of A New Tumor Gene,  
TITLE OF INVENTION: Int6, And The Use Of Reagents Derived From

TITLE OF INVENTION: These Sequences In Diagnostic Assays,  
TITLE OF INVENTION: Vaccines, Immunotherapy And Gene Therapy  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MS WORD 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/378,842  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/875,847  
FILING DATE: 09-FEB-1996  
APPLICATION NUMBER: 08/385,998  
FILING DATE: 09-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: William S. Feiler  
REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2026-4179PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 396 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
US-09-378-842-4

Query Match 100.0%; Score 396; DB 3; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MVDFAVDYKXLYSDDIPHALREKRTTVAQLKQLQAEETPIVKMFEDPETTRQMQSTRD 60  
Db 1 MVDFAVDYKXLYSDDIPHALREKRTTVAQLKQLQAEETPIVKMFEDPETTRQMQSTRD 60  
Qy 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFVRVLVPATDRLNSSL 120  
Db 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFVRVLVPATDRLNSSL 120  
Qy 121 WGLASILMKNQWDAAMEDLTRKETIDNNSVSSPLOSQLOQTWLIHWSLVFFNHPKGR 180  
Db 121 WGLASILMKNQWDAAMEDLTRKETIDNNSVSSPLOSQLOQTWLIHWSLVFFNHPKGR 180  
Qy 181 DNIIDFLYQPYQLNAIQTMCPHILRYLTAVITNKDVRKRRQVLKOLVKVIOQESYTK 240  
Db 181 DNIIDFLYQPYQLNAIQTMCPHILRYLTAVITNKDVRKRRQVLKOLVKVIOQESYTK 240  
Qy 241 DPITEFVECLYVNFDFDGAQKKLRECESLVNDFFLVACLEDFTENARLPIFETFCRIHQ 300  
Db 241 DPITEFVECLYVNFDFDGAQKKLRECESLVNDFFLVACLEDFTENARLPIFETFCRIHQ 300  
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360  
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360  
Qy 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396  
Db 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396

```
RESULT 5
US-09-858-152B-2
; Sequence 2, Application US/09858152B
; Patent No. 6737251
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
; CURRENT APPLICATION NUMBER: US/09/858,152B
; PRIOR FILING DATE: 2001-05-14
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Murine INT6
US-09-858-152B-2

Query Match 100.0%; Score 396; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDPMDDVYKLYSDDDIPHALREKRTTVAQLKQLAETEPVVKMFEDPETTRQMSTRD 60
Db 1 MVDPMDDVYKLYSDDDIPHALREKRTTVAQLKQLAETEPVVKMFEDPETTRQMSTRD 60
Qy 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
Db 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
Qy 121 WGLASEILMQNWDAAEDLTRKETIDNNSVSPLOSQLOQTWLIHWSLFFVFNHPKGR 180
Db 121 WGLASEILMQNWDAAEDLTRKETIDNNSVSPLOSQLOQTWLIHWSLFFVFNHPKGR 180
Qy 181 DNIIDFLYQPYLNAIQTWCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240
Db 181 DNIIDFLYQPYLNAIQTWCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLWVDFLVACLEDFTENARLFIETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLWVDFLVACLEDFTENARLFIETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
Qy 361 SLFSRQMLAMNTEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLFSRQMLAMNTEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 6
US-09-858-152B-4
; Sequence 4, Application US/09858152B
; Patent No. 6737251
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
```

```
; CURRENT APPLICATION NUMBER: US/09/858,152B
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 09/858,152
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 4
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-152B-4

Query Match 100.0%; Score 396; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDPMDDVYKLYSDDDIPHALREKRTTVAQLKQLAETEPVVKMFEDPETTRQMSTRD 60
Db 1 MVDPMDDVYKLYSDDDIPHALREKRTTVAQLKQLAETEPVVKMFEDPETTRQMSTRD 60
Qy 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
Db 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
Qy 121 WGLASEILMQNWDAAEDLTRKETIDNNSVSPLOSQLOQTWLIHWSLFFVFNHPKGR 180
Db 121 WGLASEILMQNWDAAEDLTRKETIDNNSVSPLOSQLOQTWLIHWSLFFVFNHPKGR 180
Qy 181 DNIIDFLYQPYLNAIQTWCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240
Db 181 DNIIDFLYQPYLNAIQTWCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLWVDFLVACLEDFTENARLFIETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLWVDFLVACLEDFTENARLFIETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
Qy 361 SLFSRQMLAMNTEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLFSRQMLAMNTEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 7
US-09-976-594-841
; Sequence 841, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 841
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 1330056CD1
US-09-976-594-841

Query Match 100.0%; Score 396; DB 4; Length 445;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MVDPMDDVYKNLYSDDIPHALREKRTTVAQLKOLQAEETPIVKMPEDPETTRMQSTRD 60  
DB 50 MVDPMDDVYKNLYSDDIPHALREKRTTVAQLKOLQAEETPIVKMPEDPETTRMQSTRD 109  
QY 61 GRMLFDYLDKKGFRQBYDLTLYRYAKFQYECGNYSGAAEYLYFPRVLVPATDRLNALSSL 120  
DB 110 GRMLFDYLDKKGFRQBYDLTLYRYAKFQYECGNYSGAAEYLYFPRVLVPATDRLNALSSL 169  
QY 121 WGLKASILQNDWDAAMEDLTRLKETIDNNSSVSPLOSLOQRTWLIIHWSLFPVFNHPKGR 180  
DB 170 WGLKASILQNDWDAAMEDLTRLKETIDNNSSVSPLOSLOQRTWLIIHWSLFPVFNHPKGR 229  
QY 181 DNIIDFLYQOYLNAIQTCWCPHLRYLTAVITNKDVKRQVLDLVKVIQOESYTYK 240  
DB 230 DNIIDFLYQOYLNAIQTCWCPHLRYLTAVITNKDVKRQVLDLVKVIQOESYTYK 289  
QY 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTIENARLFIETFCRIHQ 300  
DB 290 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTIENARLFIETFCRIHQ 349  
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360  
DB 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 409  
QY 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTODSGFY 396  
DB 410 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTODSGFY 445

## RESULT 8

US-09-270-767-46143  
; Sequence 46143, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 46143  
; LENGTH: 53  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-46143

Query Match 1.8%; Score 7; DB 4; Length 53;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 QSTRDGR 62

DB 39 QSTRDGR 45

## RESULT 9

US-09-270-767-35620  
; Sequence 35620, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 35620  
; LENGTH: 55  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-35620

Query Match 1.8%; Score 7; DB 4; Length 55;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 ALSSLWG 122

DB 47 ALSSLWG 53

## RESULT 10

US-09-270-767-50837  
; Sequence 50837, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 50837  
; LENGTH: 55  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-50837

Query Match 1.8%; Score 7; DB 4; Length 55;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 ALSSLWG 122

DB 47 ALSSLWG 53

## RESULT 11

US-09-621-976-7296  
; Sequence 7296, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano J.V.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 7296  
; LENGTH: 62  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 62  
; OTHER INFORMATION: Xaa = \*, Gly  
US-09-621-976-7296

Query Match 1.8%; Score 7; DB 4; Length 62;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 RNALSSL 120

DB 7 RNALSSL 13

```
RESULT 12
US-09-621-976-4933
; Sequence 4933, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4933
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4933

Query Match 1.8%; Score 7; DB 4; Length 70;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 114 RNALSSL 120
Db 7 RNALSSL 13

RESULT 13
US-09-025-151-8
; Sequence 8, Application US/09025151
; Patent No. 6187535
; GENERAL INFORMATION:
; APPLICANT: Legrain, Pierre
; APPLICANT: Fromont, Micheline
; APPLICANT: Rain, Jean-Christophe
; TITLE OF INVENTION: FAST AND EXHAUSTIVE METHOD FOR SELECTING A PREY
; TITLE OF INVENTION: POLYPEPTIDE INTERACTING WITH A BAIT POLYPEPTIDE OF
; TITLE OF INVENTION: INTEREST: APPLICATION TO THE CONSTRUCTION OF MAPS OF
; TITLE OF INVENTION: INTERACTORS POLYPEPTIDES
; FILE REFERENCE: 03495-0164
; CURRENT APPLICATION NUMBER: US/09/025,151
; CURRENT FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-025-151-8

Query Match 1.8%; Score 7; DB 3; Length 74;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 RDGRMLF 65
Db 14 RDGRMLF 20

RESULT 14
US-09-637-240-8
; Sequence 8, Application US/09637240
; Patent No. 6531284
; GENERAL INFORMATION:
; APPLICANT: Legrain, Pierre
; APPLICANT: Fromont, Micheline
; APPLICANT: Rain, Jean-Christophe
; TITLE OF INVENTION: FAST AND EXHAUSTIVE METHOD FOR SELECTING A PREY
; TITLE OF INVENTION: POLYPEPTIDE INTERACTING WITH A BAIT POLYPEPTIDE OF
; TITLE OF INVENTION: INTEREST: APPLICATION TO THE CONSTRUCTION OF MAPS OF
```

---

```
; TITLE OF INVENTION: INTERACTORS POLYPEPTIDES
; FILE REFERENCE: 03495-0164
; CURRENT APPLICATION NUMBER: US/09/637,240
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/025,151
; PRIOR FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-637-240-8

Query Match 1.8%; Score 7; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 RDGRMLF 65
Db 14 RDGRMLF 20

RESULT 15
US-09-621-976-7295
; Sequence 7295, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7295
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7295

Query Match 1.8%; Score 7; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 114 RNALSSL 120
Db 7 RNALSSL 13

Search completed: March 8, 2005, 20:23:34
Job time : 45 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 19:47:06 ; Search time 164 Seconds  
(without alignments)  
933.885 Million cell updates/sec

Title: US-10-783-415-4

Perfect score: 2071

Sequence: 1 MVDPMYVKNLYSDIDPHA.....LNQSRSEAPNWTQSGFY 396

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A: Geneseq 16Dec04: \*

1: Geneseqp1980s: \*

2: Geneseqp1990s: \*

3: Geneseqp2000s: \*

4: Geneseqp2001s: \*

5: Geneseqp2002s: \*

6: Geneseqp2003as: \*

7: Geneseqp2003bs: \*

8: Geneseqp2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2071	100.0	396	2	Aaw02113 Murine In
2	2071	100.0	396	2	Aaw02112 Human hom
3	2071	100.0	396	2	Aab47920 Murine In
4	2071	100.0	396	2	Aab47921 Human Int
5	2071	100.0	396	6	Abu05167 Human exp
6	2071	100.0	396	6	Abu05169 Human exp
7	2071	100.0	396	6	Abu05158 Human exp
8	2071	100.0	396	7	Adj70258 Human hea
9	2071	100.0	396	8	Adn30996 Human Int
10	2071	100.0	396	8	Adn30994 Murine In
11	2071	100.0	445	5	Abb57306 Mouse isc
12	2071	100.0	445	6	Abn39935 Human pro
13	2071	100.0	445	6	Abu05160 Human exp
14	2071	100.0	445	6	Abu05163 Human exp
15	2071	100.0	445	6	Abu05157 Human exp
16	2071	100.0	445	6	Abu05166 Human exp
17	2071	100.0	445	6	Abu05159 Human exp
18	2071	100.0	445	6	Abu05162 Human exp
19	2071	100.0	445	6	Abu05164 Human exp
20	2071	100.0	445	6	Abu05168 Human exp
21	2071	100.0	445	7	Abm85381 Human pro
22	2071	100.0	445	8	Adl13112 Human ste
23	2071	100.0	445	8	Abm81005 Tumour-as
24	2067	99.8	445	6	Abu05161 Human exp
25	2029	98.0	445	6	Abu05165 Human exp

## ALIGNMENTS

### RESULT 1

AAW02113  
ID AAW02113 standard; protein; 396 AA.

XX AAW02113;

DT 15-MAY-1997 (first entry)

DE Murine Int6 protein associated with MMTV integration and tumour growth.

KW MMTV; mouse mammary tumour virus; Int6; breast cancer; neoplasia;

KW diagnosis; treatment; immunotherapy; vaccine; probe; primer.

OS Mus musculus.

Key Location/Qualifiers

FT Misc-difference 27 /note= "potential cAMP/cGMP-dependent protein kinase phosphorylation site"

FT Misc-difference 51 /note= "potential protein kinase C phosphorylation site"

FT Misc-difference 57 /note= "potential casein kinase II phosphorylation site"

FT Misc-difference 90 /note= "potential tyrosine kinase phosphorylation site"

FT Misc-difference 94 /note= "potential glycosylation site"

FT Misc-difference 112 /note= "potential protein kinase C phosphorylation site"

FT Misc-difference 149 /note= "potential glycosylation site"

FT Misc-difference 214 /note= "potential casein kinase II phosphorylation site"

FT Misc-difference 238 /note= "potential casein kinase II phosphorylation site"

FT Misc-difference 314 /note= "potential casein kinase II phosphorylation site"

FT Misc-difference 363 /note= "potential protein kinase C phosphorylation site"

FT Misc-difference 381 /note= "potential casein kinase II phosphorylation site"

XX WO9624672-A1.

XX 15-AUG-1996.

PP 09-FEB-1996; 96WO-US001884.

XX

26	1279.5	61.8	320	7	ABM85380	Mouse pro
27	1220.5	58.9	435	4	ABE64579	Drosophil
28	1104.5	53.3	302	4	ABG00533	Novel hum
29	1104.5	53.3	302	4	ABG21243	Novel hum
30	1090.5	52.7	427	3	AAG34208	Zea mays
31	1090.5	52.7	436	3	AAG34207	Zea mays
32	1090.5	52.7	467	3	AAG34206	Zea mays
33	903	43.6	418	8	ADM57227	A thalian
34	838.5	40.5	1202	4	ABG21244	Novel hum
35	838.5	40.5	1202	7	ADF60389	Human con
36	738.5	35.7	155	4	ABG21242	Novel hum
37	564	27.2	117	5	ABP43106	Human ova
38	537.5	26.0	197	3	AAG16552	Arabidops
39	537.5	26.0	250	3	AAG16551	Arabidops
40	537.5	26.0	274	3	AAG16550	Arabidops
41	223	10.8	45	4	AAM22004	Peptide #
42	223	10.8	45	4	ABB44381	Peptide #
43	223	10.8	45	4	AAM38359	Peptide #
44	223	10.8	45	4	ABB27240	Protein #
45	223	10.8	45	4	AAM78125	Human bon

```

PR 09-FEB-1995; 95US-00385998.
XX
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI Marchetti A, Buttitta F, Smith GH, Callahan R;
XX
XX DR WPI; 1996-384444/38.
XX
XX DR N-PSDB; AAT36177.
XX
XX PT DNA encoding Int6 tumour associated protein - and use of reagents derived
XX from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.
XX
XX PS Disclosure; Page 60-61; 93pp; English.
XX
XX CC AAW02113 is the murine Int6 protein. The Int6 gene is located at
XX chromosome 15 of a mouse genome. The Int6 gene is associated with MMTV
XX (mouse mammary tumour virus) integration into a host genome during
XX tumorigenesis. Primers and probes may be derived from the Int6 gene
XX sequence and used for detection in assays to diagnose MMTV infection, or
XX any other Int6 gene integration. Antibodies against the Int6 protein can
XX be used in the same way. The DNA and protein may also be used gene
XX therapy and vaccines
XX
XX SQ Sequence 396 AA;
Query Match 100.0%; Score 2071; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFAVDVYKNLYSDDI PHALREKRTTVAQLKQLQAEITEPIVKMFEDPETTRQMOSTRD 60
DB 1 MVDFAVDVYKNLYSDDI PHALREKRTTVAQLKQLQAEITEPIVKMFEDPETTRQMOSTRD 60
QY 61 GRMLFDYLADKHGFRQEYLDLTLYRYAKFQYECGNYSGAAEVLFFRVLVPATDRNALSSL 120
DB 61 GRMLFDYLADKHGFRQEYLDLTLYRYAKFQYECGNYSGAAEVLFFRVLVPATDRNALSSL 120
QY 121 WGKLASEILMQNWDAAMEDLTRLKETIDNNSVSSPLQSLQORTWLIHWSLFFVFNHPKGR 180
DB 121 WGKLASEILMQNWDAAMEDLTRLKETIDNNSVSSPLQSLQORTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240
DB 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWIVNLRNARDAKIDSKLGHVVMGNNVSPYQQVIEKTK 360
DB 301 CISINMLADKLNMTPEEAERWIVNLRNARDAKIDSKLGHVVMGNNVSPYQQVIEKTK 360
QY 361 SLSPFSQMLAMNIEKKLNQSRSEAPNATQDSGFY 396
DB 361 SLSPFSQMLAMNIEKKLNQSRSEAPNATQDSGFY 396
RESULT 2
AAW02112
ID AAW02112 standard; protein; 396 AA.
XX
XX AC AAW02112;
XX
XX DT 14-MAY-1997 (first entry)
XX
XX DE Human homologue of Int6 protein associated with MMTV integration.
XX
XX KW MMTV; mouse mammary tumour virus; Int6; breast cancer; neoplasia;
XX KW diagnosis; treatment; immunotherapy; vaccine; probe; primer.
XX
XX OS Homo sapiens.
XX

```

```

PN WO9624672-A1.
XX
XX PD 15-AUG-1996.
XX
XX PF 09-FEB-1996; 96WO-US001884.
XX
XX PR 09-FEB-1995; 95US-00385998.
XX
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI Marchetti A, Buttitta F, Smith GH, Callahan R;
XX
XX DR WPI; 1996-384444/38.
XX
XX DR N-PSDB; AAT36148.
XX
XX PT DNA encoding Int6 tumour associated protein - and use of reagents derived
XX from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.
XX
XX PS Claim 3; Page 62-63; 93pp; English.
XX
XX CC AAW02112 represents the human homologue of murine Int6 protein. The Int6
XX gene is located at chromosome 15 of the mouse genome and is associated
XX with MMTV (mouse mammary tumour virus) integration into a host genome
XX during tumorigenesis. Primers and probes may be derived from the Int6
XX cDNA sequence and used for the detection of the Int6 gene. These can be
XX used in assays to diagnose MMTV infection, or any other Int6 gene
XX integration. Antibodies against the Int6 protein can be used in the same
XX way
XX
XX SQ Sequence 396 AA;
Query Match 100.0%; Score 2071; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFAVDVYKNLYSDDI PHALREKRTTVAQLKQLQAEITEPIVKMFEDPETTRQMOSTRD 60
DB 1 MVDFAVDVYKNLYSDDI PHALREKRTTVAQLKQLQAEITEPIVKMFEDPETTRQMOSTRD 60
QY 61 GRMLFDYLADKHGFRQEYLDLTLYRYAKFQYECGNYSGAAEVLFFRVLVPATDRNALSSL 120
DB 61 GRMLFDYLADKHGFRQEYLDLTLYRYAKFQYECGNYSGAAEVLFFRVLVPATDRNALSSL 120
QY 121 WGKLASEILMQNWDAAMEDLTRLKETIDNNSVSSPLQSLQORTWLIHWSLFFVFNHPKGR 180
DB 121 WGKLASEILMQNWDAAMEDLTRLKETIDNNSVSSPLQSLQORTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240
DB 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWIVNLRNARDAKIDSKLGHVVMGNNVSPYQQVIEKTK 360
DB 301 CISINMLADKLNMTPEEAERWIVNLRNARDAKIDSKLGHVVMGNNVSPYQQVIEKTK 360
QY 361 SLSPFSQMLAMNIEKKLNQSRSEAPNATQDSGFY 396
DB 361 SLSPFSQMLAMNIEKKLNQSRSEAPNATQDSGFY 396
RESULT 3
AAW47920
ID AAW47920 standard; protein; 396 AA.
XX
XX AC AAW47920;
XX
XX DT 21-MAY-2002 (first entry)
XX
XX DE Murine Int6.

```

XX Murine; human; Int6; integration site; deregulation; neoplasia;  
 KW mouse mammary tumour virus; MMV; cancer; immunotherapy; gene therapy;  
 KW prenatal screening; foetus; vaccine.  
 XX Mus musculus.

XX Key Location/Qualifiers  
 XX PH 27 /note= "Phosphorylation site for cAMP/cGMP-dependent  
 FT Modified-site protein kinase"  
 FT 51  
 FT Modified-site /note= "Phosphorylation site for protein kinase C"  
 FT 57  
 FT Modified-site /note= "Phosphorylation site for casein kinase II"  
 FT 90  
 FT Modified-site /note= "Phosphorylation site for tyrosine kinase"  
 FT 94  
 FT Modified-site /note= "Phosphorylation site for tyrosine kinase and  
 FT glycosylation site"  
 FT 112  
 FT Modified-site /note= "Phosphorylation site for protein kinase C"  
 FT 149  
 FT Modified-site /note= "Phosphorylation site for tyrosine kinase and  
 FT glycosylation site"  
 FT 214  
 FT Modified-site /note= "Phosphorylation site for casein kinase II"  
 FT 238  
 FT Modified-site /note= "Phosphorylation site for casein kinase II"  
 FT 314  
 FT Modified-site /note= "Phosphorylation site for casein kinase II"  
 FT 363  
 FT Modified-site /note= "Phosphorylation site for protein kinase C"  
 FT 381  
 FT Modified-site /note= "Phosphorylation site for casein kinase II"

XX US6342392-B1.  
 XX PN 29-JAN-2002.  
 XX PD 23-AUG-1999; 99US-00378842.  
 XX PF 09-FEB-1995; 95US-00385998.  
 XX PR 09-FEB-1996; 96WO-US001884.  
 XX PR 25-SEP-1997; 97US-00875847.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX PA Marchetti A, Buttitta F, Smith GH, Callahan R;  
 XX PI WPI; 1996-384444/38.  
 XX DR N-PSDB; AAI72498.  
 XX DNA encoding Int6 tumour associated protein - and use of reagents derived  
 PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.  
 XX Disclosure; Col 35-38; 45pp; English.

XX This sequence shows murine Int6. Int6 is an integration site for mouse  
 CC mammary tumour virus (MMTV), which causes deregulation of expression of  
 CC cellular genes adjacent to the site of MMTV integration in mammary  
 CC tumours. The Int6 protein has been found to be highly conserved across  
 CC species, with Drosophila Int6 being 60% identical to human/mouse Int6.  
 CC This indicates that Int6 is serving a basic life function. The method of  
 CC the invention comprises assaying a sample to detect a human Int6 nucleic  
 CC acid sequence, or its fragment, by contacting the sample with a sequence  
 CC of at least 15 consecutive nucleotides of human Int6 cDNA or a  
 CC conservative variant of it, where a disrupted expression or loss of  
 CC expression of the variant is associated with neoplasia. The method is  
 CC useful for prenatal screening of a foetus or to pre-symptomatically  
 CC screen a subject at risk of having cancer. Detecting mutations in the  
 CC Int6 gene can provide diagnostic and prognostic information. The nucleic  
 CC acids and proteins are useful in immunotherapy, gene therapy or as

CC vaccines for treating or preventing cancer. The nucleic acids are useful  
 CC as probes for isolating homologues of Int6 gene or for detecting  
 CC mutations in the Int6 gene  
 XX SQ Sequence 396 AA;  
 Query Match 100.0%; Score 2071; DB 2; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 7,2e-193;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MYDFAMDVYKXLYSDDIPHALREKRTTVVAQLKQLQAEPIVKMFPEDPETTRQMOSTRD 60  
 DB 1 MYDFAMDVYKXLYSDDIPHALREKRTTVVAQLKQLQAEPIVKMFPEDPETTRQMOSTRD 60  
 QY 61 GRMLFDYLDKRGHGFQREYLDTLRYAKFYQECGNYSGAAEYLYFFRVLPATDRNALSSL 120  
 DB 61 GRMLFDYLDKRGHGFQREYLDTLRYAKFYQECGNYSGAAEYLYFFRVLPATDRNALSSL 120  
 QY 121 WGLKASEILMQNWDAMEDLTRKETIDNNSVSSPLQSLQQRRTWLHWSLFFVFNHPKGR 180  
 DB 121 WGLKASEILMQNWDAMEDLTRKETIDNNSVSSPLQSLQQRRTWLHWSLFFVFNHPKGR 180  
 QY 181 DNIIDFLYQPOYLNAIOTMCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240  
 DB 181 DNIIDFLYQPOYLNAIOTMCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240  
 QY 241 DPITFEVCELYNFDGQAOKLRCESVLVNDFLVACLEDPIENARLFIETFCRIHQ 300  
 DB 241 DPITFEVCELYNFDGQAOKLRCESVLVNDFLVACLEDPIENARLFIETFCRIHQ 300  
 QY 301 CUSINMLADKLNMTPEEAERWIVNLI RNARLDAKIDSKLGHVVMGNNAVSPYQOQVIEKTK 360  
 DB 301 CUSINMLADKLNMTPEEAERWIVNLI RNARLDAKIDSKLGHVVMGNNAVSPYQOQVIEKTK 360  
 QY 361 SLSPFSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396  
 DB 361 SLSPFSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396

RESULT 4  
 AAB47921  
 ID AAB47921 standard; protein; 396 AA.  
 AC AAB47921;  
 XX 21-MAY-2002 (first entry)  
 DE Human Int6.  
 KW Murine; human; Int6; integration site; deregulation; neoplasia;  
 KW mouse mammary tumour virus; MMV; cancer; immunotherapy; gene therapy;  
 KW prenatal screening; foetus; vaccine; chromosome 8q22-Q24.  
 OS Homo sapiens.  
 XX US6342392-B1.  
 XX 29-JAN-2002.  
 XX 23-AUG-1999; 99US-00378842.  
 XX 09-FEB-1995; 95US-00385998.  
 XX 09-FEB-1996; 96WO-US001884.  
 XX 25-SEP-1997; 97US-00875847.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX PI Marchetti A, Buttitta F, Smith GH, Callahan R;  
 XX WPI; 1996-384444/38.  
 XX DR N-PSDB; AAI72498.  
 XX DNA encoding Int6 tumour associated protein - and use of reagents derived

PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.  
PS Disclosure; Col 39-44; 45pp; English.

XX This sequence shows human Int6. The human Int6 coding sequence was  
CC isolated by using murine Int6 sequences as probes. Human Int6 is  
CC organised into 13 exons as is murine Int6, and contains a CA-repeat in  
CC the 7th intron. Human Int6 has been localised to chromosome 8, more  
CC specifically to 8q22-q24. Int6 is an integration site for mouse mammary  
CC tumour virus (MMTV), which causes deregulation of expression of cellular  
CC genes adjacent to the site of MMTV integration in mammary tumours. The  
CC Int6 protein has been found to be highly conserved across species, with  
CC Drosophila Int6 being 60% identical to human/mouse Int6. This indicates  
CC that Int6 is serving a basic life function. The method of the invention  
CC comprises assaying a sample to detect a human Int6 nucleic acid sequence,  
CC or its fragment, by contacting the sample with a sequence of at least 15  
CC consecutive nucleotides of human Int6 cDNA or a conservative variant of  
CC it, where a disrupted expression or loss of expression of the variant is  
CC associated with neoplasia. The method is useful for prenatal screening of  
CC a foetus or to pre-symptomatically screen a subject at risk of having  
CC cancer. Detecting mutations in the Int6 gene can provide diagnostic and  
CC prognostic information. The nucleic acids and proteins are useful in  
CC immunotherapy, gene therapy or as vaccines for treating or preventing  
CC cancer. The nucleic acids are useful as probes for isolating homologues  
CC of Int6 gene or for detecting mutations in the Int6 gene

XX Sequence 396 AA;

Query Match 100.0%; Score 2071; DB 2; Length 396;  
Best Local Similarity 100.0%; Pred. No. 7.2e-193;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVDFAVDVKNLYSDDIIPHALREKRTTVVAOLKQIAETEPVKMFEDPETTRQMOSTRD 60  
DB 1 MVDFAVDVKNLYSDDIIPHALREKRTTVVAOLKQIAETEPVKMFEDPETTRQMOSTRD 60  
QY 61 GRMLFDYLADKHGFRQVEYLDLTYRYAKFQYECGNYSGAAEYLFPFVLVPATDRNALSSL 120  
DB 61 GRMLFDYLADKHGFRQVEYLDLTYRYAKFQYECGNYSGAAEYLFPFVLVPATDRNALSSL 120  
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSQLOORTWLIHWSLVFFNHPKGR 180  
DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSQLOORTWLIHWSLVFFNHPKGR 180  
QY 181 DNIIDFLYQPYQLNAIOTMCPHILRYLTAVITNKVKRQVLDKLVKVIQOESYTYK 240  
DB 181 DNIIDFLYQPYQLNAIOTMCPHILRYLTAVITNKVKRQVLDKLVKVIQOESYTYK 240  
QY 241 DPITEFVECLYVNFDFDGAQKKLRCEESVLVNDFLVACLEDFIENARLFIFETFCRIHQ 300  
DB 241 DPITEFVECLYVNFDFDGAQKKLRCEESVLVNDFLVACLEDFIENARLFIFETFCRIHQ 300  
QY 301 CISINMLADKLNMTPEERIVNLIARNLDAKIDSKLGHVVMGNNAVSPYQVIEKTK 360  
DB 301 CISINMLADKLNMTPEERIVNLIARNLDAKIDSKLGHVVMGNNAVSPYQVIEKTK 360  
QY 361 SLSFRSOMLANNIEKKLNQNSSEAPNATQDSGFY 396  
DB 361 SLSFRSOMLANNIEKKLNQNSSEAPNATQDSGFY 396

RESULT 5

ABU05167

ID ABU05167 standard; protein; 396 AA.

XX AC

XX ABU05167;

XX DT 29-JAN-2003 (first entry)

XX DE Human expressed protein tag (EPT) #1833.

XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;

KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX OS Homo sapiens.

XX WO200278524-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US009671.

XX PR 28-MAR-2001; 2001US-0279495P.

XX PR 21-MAY-2001; 2001US-0292544P.

XX PR 08-AUG-2001; 2001US-0310801P.

XX PR 01-OCT-2001; 2001US-0326370P.

XX PR 04-DEC-2001; 2001US-0336780P.

XX PR 20-FEB-2002; 2002US-0358985P.

XX PA (ZYCO-) ZYCOS INC.

XX PI Chicx RM, Tomlinson AJ, Urban RG;

XX DR WPI; 2003-040607/03.

XX PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX Example 2; SEQ ID NO 1833; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 396 AA;

Query Match 100.0%; Score 2071; DB 6; Length 396;  
Best Local Similarity 100.0%; Pred. No. 7.2e-193;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVKNLYSDDIIPHALREKRTTVVAOLKQIAETEPVKMFEDPETTRQMOSTRD 60  
DB 1 MVDFAVDVKNLYSDDIIPHALREKRTTVVAOLKQIAETEPVKMFEDPETTRQMOSTRD 60

QY 61 GRMLFDYLADKHGFRQVEYLDLTYRYAKFQYECGNYSGAAEYLFPFVLVPATDRNALSSL 120  
DB 61 GRMLFDYLADKHGFRQVEYLDLTYRYAKFQYECGNYSGAAEYLFPFVLVPATDRNALSSL 120

QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSQLOORTWLIHWSLVFFNHPKGR 180  
DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSQLOORTWLIHWSLVFFNHPKGR 180

QY 181 DNIIDFLYQPYQLNAIOTMCPHILRYLTAVITNKVKRQVLDKLVKVIQOESYTYK 240  
DB 181 DNIIDFLYQPYQLNAIOTMCPHILRYLTAVITNKVKRQVLDKLVKVIQOESYTYK 240

QY 241 DPITEFVECLYVNFDFDGAQKKLRCEESVLVNDFLVACLEDFIENARLFIFETFCRIHQ 300



Db 241 DPTFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300  
QY 301 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQQVIEKTK 360  
Db 301 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQQVIEKTK 360  
QY 361 SLSPRSQMLAMNIEKLNQNSRSEAPNWTQDSGFY 396  
Db 361 SLSPRSQMLAMNIEKLNQNSRSEAPNWTQDSGFY 396

## RESULT 6

ABU05169  
ID ABU05169 standard; protein; 396 AA.

XX AC ABU05169;

XX DT 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1835.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX OS Homo sapiens.

XX PN WO200278524-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US009671.

XX PR 28-MAR-2001; 2001US-0279495P.

XX PR 21-MAY-2001; 2001US-0292544P.

XX PR 08-AUG-2001; 2001US-0310801P.

XX PR 01-OCT-2001; 2001US-0326370P.

XX PR 04-DEC-2001; 2001US-0336780P.

XX PR 20-FEB-2002; 2002US-0358985P.

XX PA (ZYCO-) ZYCOS INC.

XX PI Chicx RM, Tomlinson AJ, Urban RG;

XX DR WPI; 2003-040607/03.

XX PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.

XX PS Example 2; SEQ ID NO 1835; 134pp; English.

XX CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 396 AA;

Query Match 100.0%; Score 2071; DB 6; Length 396;  
Best Local Similarity 100.0%; Pred. No. 7.2e-193;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDFAMDVYKNLYSDDI PHALREKETTVAQKQLQAEETEPVVKMFDPETTRQMOSTRD 60  
Db 1 MYDFAMDVYKNLYSDDI PHALREKETTVAQKQLQAEETEPVVKMFDPETTRQMOSTRD 60

QY 61 GRMLFDYLADKHGFRQEVLDITLYRYAKFOYECGNTSGAAEYLYFFRVLVLPATDRNALSSL 120

Db 61 GRMLFDYLADKHGFRQEVLDITLYRYAKFOYECGNTSGAAEYLYFFRVLVLPATDRNALSSL 120

QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQQRITLIHWSLFFVFNHPKGR 180

Db 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQQRITLIHWSLFFVFNHPKGR 180

QY 181 DNIIDFLYQOYLNAIOTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240

Db 181 DNIIDFLYQOYLNAIOTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240

QY 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300

Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300

QY 301 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQQVIEKTK 360

Db 301 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQQVIEKTK 360

QY 361 SLSPRSQMLAMNIEKLNQNSRSEAPNWTQDSGFY 396

Db 361 SLSPRSQMLAMNIEKLNQNSRSEAPNWTQDSGFY 396

## RESULT 7

ABU05158

ID ABU05158 standard; protein; 396 AA.

XX AC ABU05158;

XX DT 29-JAN-2003 (first entry)

XX DE Human expressed protein tag (EPT) #1824.

XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX OS Homo sapiens.

XX PN WO200278524-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US009671.

XX PR 28-MAR-2001; 2001US-0279495P.

XX PR 21-MAY-2001; 2001US-0292544P.

XX PR 08-AUG-2001; 2001US-0310801P.

XX PR 01-OCT-2001; 2001US-0326370P.

XX PR 04-DEC-2001; 2001US-0336780P.

XX PR 20-FEB-2002; 2002US-0358985P.

XX PA (ZYCO-) ZYCOS INC.

XX PI Chicx RM, Tomlinson AJ, Urban RG;

XX DR WPI; 2003-040607/03.

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX  
PS Example 2; SEQ ID NO 1824; 134pp; English.  
XX  
CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 396 AA;

Query Match 100.0%; Score 2071; DB 6; Length 396;  
Best Local Similarity 100.0%; Pred. No. 7.2e-193;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVDFAVDYKLYSDSDIPHALREKRTTVVAQLKQLQAEPIVKMFEDPETTRMQSTRD 60  
Db 1 MVDFAVDYKLYSDSDIPHALREKRTTVVAQLKQLQAEPIVKMFEDPETTRMQSTRD 60  
QY 61 GRMLFDYLDADKHGFRQYEDLTLYRYAKFYQYECGNYSGAAEYLYFPRVLVPATDRNALSSL 120  
Db 61 GRMLFDYLDADKHGFRQYEDLTLYRYAKFYQYECGNYSGAAEYLYFPRVLVPATDRNALSSL 120  
QY 121 WGLKASEILMNQWDAAMEDLTLKKTETIDNNSSVSPLOSLOQTWLIHWSLVFFFNHPKGR 180  
Db 121 WGLKASEILMNQWDAAMEDLTLKKTETIDNNSSVSPLOSLOQTWLIHWSLVFFFNHPKGR 180  
QY 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDKVRKQVLDKLVKVIQBSYTYK 240  
Db 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDKVRKQVLDKLVKVIQBSYTYK 240  
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300  
Db 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300  
QY 301 CISINMLADKLMNTPTEAEERWTNLIARNLDKIDSKLGHVGMGNNAVSPYQVIEKTK 360  
Db 301 CISINMLADKLMNTPTEAEERWTNLIARNLDKIDSKLGHVGMGNNAVSPYQVIEKTK 360  
QY 361 SLSPRSQMLAMNIEKLNQNSRSEAPNATQDSGFY 396  
Db 361 SLSPRSQMLAMNIEKLNQNSRSEAPNATQDSGFY 396

RESULT 8  
ADJ70258  
ID ADJ70258 standard; protein; 396 AA.  
XX  
AC ADJ70258;  
XX  
DT 06-MAY-2004 (first entry)  
DE Human heat mitochondrial protein as a therapeutic target SeqID2064.  
XX  
KW mitochondrial; human; screening assay; diabetes mellitus;  
KW Huntington's disease; osteoarthritis;  
KW Leber's hereditary optic neuropathy; LHON;

KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;  
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;  
KW osteopathic; ophthalmological; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO2003087768-A2.  
XX  
PD 23-OCT-2003.  
XX  
PF 04-APR-2003; 2003WO-US010870.  
XX  
PR 12-APR-2002; 2002US-0372843P.  
PR 17-JUN-2002; 2002US-0389987P.  
PR 20-SEP-2002; 2002US-0412418P.  
XX  
XX (MITO-) MITOKOR.  
PA (BUCK-) BUCK INST AGE RES.  
PA  
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;  
PI Warnock DE;  
XX  
XX WPI; 2003-845369/78.  
XX  
PT Identifying a mitochondrial target for drug screening assays and for  
PT treating diseases associated with altered mitochondrial function,  
PT comprises detecting a modified polypeptide in a sample and correlating  
PT with the disease.  
XX  
PS Claim 1; SEQ ID NO 2064; 180pp; English.  
XX  
CC This invention relates to novel mitochondrial targets that can be used  
CC for therapeutic intervention in treating a disease associated with  
CC altered mitochondrial function. Specifically, it refers to a method for  
CC identifying proteins of the human heart mitochondrial proteome that are  
CC useful for drug screening assays, as well as therapeutic targets. The  
CC present invention describes a method for identifying such proteins that  
CC can be used in the treatment of various diseases associated with altered  
CC mitochondrial function including diabetes mellitus, Huntington's disease,  
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
CC compositions have neuroprotective, nootropic, antidiabetic,  
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
CC cytostatic activities. This polypeptide sequence is a human heart  
CC mitochondrial protein of the invention.  
XX  
SQ Sequence 396 AA;  
Query Match 100.0%; Score 2071; DB 7; Length 396;  
Best Local Similarity 100.0%; Pred. No. 7.2e-193;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVDFAVDYKLYSDSDIPHALREKRTTVVAQLKQLQAEPIVKMFEDPETTRMQSTRD 60  
Db 1 MVDFAVDYKLYSDSDIPHALREKRTTVVAQLKQLQAEPIVKMFEDPETTRMQSTRD 60  
QY 61 GRMLFDYLDADKHGFRQYEDLTLYRYAKFYQYECGNYSGAAEYLYFPRVLVPATDRNALSSL 120  
Db 61 GRMLFDYLDADKHGFRQYEDLTLYRYAKFYQYECGNYSGAAEYLYFPRVLVPATDRNALSSL 120  
QY 121 WGLKASEILMNQWDAAMEDLTLKKTETIDNNSSVSPLOSLOQTWLIHWSLVFFFNHPKGR 180  
Db 121 WGLKASEILMNQWDAAMEDLTLKKTETIDNNSSVSPLOSLOQTWLIHWSLVFFFNHPKGR 180  
QY 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDKVRKQVLDKLVKVIQBSYTYK 240  
Db 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDKVRKQVLDKLVKVIQBSYTYK 240  
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300  
Db 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300

---

Qy	181	DNIIDLFLYQPOYLNAIQTMCPHILRYLTTAVITNKDVRKERRQVLKDLVKVIOQESYTYK	240
Db	181	DNIIDLFLYQPOYLNAIQTMCPHILRYLTTAVITNKDVRKERRQVLKDLVKVIOQESYTYK	240
Qy	241	DPITEFVECLYVNFDFDGAQKKLRECESVLVNDFFLVACLEDFTIENARLFIPTFCRIHQ	300
Db	241	DPITEFVECLYVNFDFDGAQKKLRECESVLVNDFFLVACLEDFTIENARLFIPTFCRIHQ	300
Qy	301	CISINMLADKLNTPEEARWTVNLIRNARLDAKIDSKLGHVYMGNNNAVSPYQOVIETKT	360
Db	301	CISINMLADKLNTPEEARWTVNLIRNARLDAKIDSKLGHVYMGNNNAVSPYQOVIETKT	360
Qy	361	SUSFRSOMLAMNIEKKLNONSRSEAPNWTQDSGFY	396
Db	361	SUSFRSOMLAMNIEKKLNONSRSEAPNWTQDSGFY	396

RESULT 11	
ABBS57306	
ID	ABBS57306 standard; protein; 445 AA.
XX	
AC	ABBS57306;
XX	
DT	07-MAR-2002 (first entry)
XX	
DE	Mouse ischaemic condition related protein sequence SEQ ID NO:856.
XX	
KW	Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW	vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX	
OS	Mus musculus.
XX	
PN	WO200188188-A2.
XX	
PD	22-NOV-2001.
XX	
PF	18-MAY-2001; 2001WO-JP004192.
XX	
PR	18-MAY-2000; 2000JP-00145977.
XX	
XX	
PA	(UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX	
PI	Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX	
WPI	2002-034733/04.
DR	N-PSDB: ABI99770.
DR	

Examining the ischemic condition (e.g. occlusive ischemia) by measuring expression levels of particular genes defined in the specification or by determining the expression profile of a gene group comprising these genes.

Claim 2: Page 2125-2127; 2690pp; English.

The present invention describes a method for examining ischaemic conditions, comprising measuring the expression levels of particular genes (I) in a test sample or determining the expression profile of a gene group in the sample comprising genes selected from (I). The method is useful for examining the ischaemic condition (e.g. compressive ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring expression levels of particular genes (ABI99202 to ABI99912, encoding the protein sequences in AB957020 to AB957374) or by determining the expression profile of a gene group comprising these genes. The expression levels or expression profiles produced by these genes are used as an indicator when screening for ischaemic condition-improving drugs or therapeutics for ischaemic diseases. AB199913 and AB199914 represent PCR primers for a mouse ischaemic condition related sequence, which are used in the exemplification of the present invention.

Sequence 445 AA;

Query Match	Score 2071	DB 5	Length 445
100.0%			

[illegible]

RESULT 12	
ABR39935	
ID	ABR39935 standard; protein; 445 AA.
XX	
AC	ABR39935;
XX	
DT	11-AUG-2003 (first entry)
XX	
DE	Human prostate selective polypeptide Pr327.
XX	
KW	Prostate; molecular marker; cancer; cytostatic; gene therapy; human.
XX	
OS	Homo sapiens.
XX	
FN	WO2003014298-A2.
XX	
PD	20-FEB-2003.
XX	
PF	02-AUG-2002; 2002WO-US024431.
XX	
PR	03-AUG-2001; 2001US-0309470P.
PR	30-OCT-2001; 2001US-0330747P.
XX	
PA	(ORIG-) ORIGENE TECHNOLOGIES INC.
XX	
PI	Sun Z, Li X, Jay G, Kovacs KF, Shu Y, Fan W;
XX	
DR	WPI; 2003-256562/25.
DR	N-PSDB; ACC47340.
XX	
PT	New polynucleotide, useful for preparing a composition for treating
PT	prostate disease, e.g., cancer.
XX	
PS	Claim 5; Page 150-152; 212pp; English.
XX	
CC	The invention relates to prostate selective polynucleotides and
CC	polypeptides. The polynucleotides are expressed in prostate and are
CC	useful as molecular markers, as drug targets, and for detecting,
CC	monitoring, preventing or treating diseases and conditions related to
CC	prostate, such as prostate cancers. The present sequence represents a
CC	prostate specific polypeptide

XX Sequence 445 AA;  
SQ  
Query Match 100.0%; Score 2071; DB 6; Length 445;  
Best Local Similarity 100.0%; Pred. No. 8.5e-193;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MYDFAMDVYKNLYSDDI PHALREKRTTVAQKQLQAEETPIVKMPEDEPTRQMOSTRD 60  
DB 50 MYDFAMDVYKNLYSDDI PHALREKRTTVAQKQLQAEETPIVKMPEDEPTRQMOSTRD 109  
QY 61 GRMLFDYLADKHGFRQYLDTLRYAKFOYECGNYSAAEYLYFFRVLVLPATDRNALSSL 120  
DB 110 GRMLFDYLADKHGFRQYLDTLRYAKFOYECGNYSAAEYLYFFRVLVLPATDRNALSSL 169  
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 180  
DB 170 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 229  
QY 181 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQESYTYK 240  
DB 230 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQESYTYK 289  
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300  
DB 290 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 349  
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
DB 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 409  
QY 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396  
DB 410 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445  
RESULT 13  
ABU05160  
ID ABU05160 standard; protein; 445 AA.  
XX  
AC ABU05160;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1826.  
XX  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200278524-A2.  
XX  
PD 10-OCT-2002.  
XX  
PF 28-MAR-2002; 2002WO-US009671.  
XX  
PR 28-MAR-2001; 2001US-0279495P.  
PR 21-MAY-2001; 2001US-0292544P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX  
PA (ZYCO-) ZYCOS INC.  
XX  
XX Chicx RM, Tomlinson AJ, Urban RG;  
PI  
XX  
DR WPI; 2003-040607/03.  
XX

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX  
XX Example 2; SEQ ID NO 1826; 134pp; English.  
XX  
CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 445 AA;  
Query Match 100.0%; Score 2071; DB 6; Length 445;  
Best Local Similarity 100.0%; Pred. No. 8.5e-193;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MYDFAMDVYKNLYSDDI PHALREKRTTVAQKQLQAEETPIVKMPEDEPTRQMOSTRD 60  
DB 50 MYDFAMDVYKNLYSDDI PHALREKRTTVAQKQLQAEETPIVKMPEDEPTRQMOSTRD 109  
QY 61 GRMLFDYLADKHGFRQYLDTLRYAKFOYECGNYSAAEYLYFFRVLVLPATDRNALSSL 120  
DB 110 GRMLFDYLADKHGFRQYLDTLRYAKFOYECGNYSAAEYLYFFRVLVLPATDRNALSSL 169  
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 180  
DB 170 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 229  
QY 181 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQESYTYK 240  
DB 230 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQESYTYK 289  
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300  
DB 290 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 349  
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
DB 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 409  
QY 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396  
DB 410 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445  
RESULT 14  
ABU05163  
ID ABU05163 standard; protein; 445 AA.  
XX  
AC ABU05163;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1829.  
XX  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;

```
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
OS Homo sapiens.
XX
XX WO200278524-A2.
XX PD 10-OCT-2002.
XX PF 28-MAR-2002; 2002WO-US009671.
XX PR 28-MAR-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.
XX PR 01-OCT-2001; 2001US-0326370P.
XX PR 04-DEC-2001; 2001US-0336780P.
XX PR 20-FEB-2002; 2002US-0358985P.
XX PA (ZYCO-) ZYCOS INC.
XX PI Chicx RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX Example 2; SEQ ID NO 1829; 134pp; English.
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Query Match 100.0%; Score 2071; DB 6; Length 445;
XX Best Local Similarity 100.0%; Pred. No. 8.5e-193;
XX Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MVDFAVDYKNLYSDIPHALREKTTVVAQLKQLQAEPTVKMFEDPETTRQMSTRD 60
XX |||||
XX 50 MVDFAVDYKNLYSDIPHALREKTTVVAQLKQLQAEPTVKMFEDPETTRQMSTRD 109
XX |||||
XX 61 GRMLFDYLDADKHGFRQYLDLTLYRYAKFYECGNYSGAAEYLYFFRVLVLPATDRNALSSL 120
XX |||||
XX 110 GRMLFDYLDADKHGFRQYLDLTLYRYAKFYECGNYSGAAEYLYFFRVLVLPATDRNALSSL 169
XX |||||
XX 121 WGLKASEILMQWDAAMEDLTFLKETIDNNSVSSPLOSQORTWLHWSLFFVFNHPKGR 180
XX |||||
XX 170 WGLKASEILMQWDAAMEDLTFLKETIDNNSVSSPLOSQORTWLHWSLFFVFNHPKGR 229
XX |||||
XX 181 DNIIDFLYQPOLYNAIOTMCPHILRYLTAVITNKDKRQVLDKLVKVYQOESYTYK 240
XX |||||
XX 230 DNIIDFLYQPOLYNAIOTMCPHILRYLTAVITNKDKRQVLDKLVKVYQOESYTYK 289
XX |||||
XX 241 DPITEFVECLYVNFDFDGAQKLCESVLVNDFLVACLEDFIENARLFIFETFCRIHQ 300
XX |||||
```

```
Db 290 DPITEFVECLYVNFDFDGAQKLCESVLVNDFLVACLEDFIENARLFIFETFCRIHQ 349
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARDAKIDSGLGHVVMGNNNAVSPYQQVIEKTK 360
Db 350 CISINMLADKLNMTPEEAERWIVNLIRNARDAKIDSGLGHVVMGNNNAVSPYQQVIEKTK 409
Qy 361 SLSFRSQMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 410 SLSFRSQMLAWNIEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 15
ABU05157
ID ABU05157 standard; protein; 445 AA.
XX AC ABU05157;
XX DT 29-JAN-2003 (first entry)
XX DE Human expressed protein tag (EPT) #1823.
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX OS Homo sapiens.
XX PN WO200278524-A2.
XX PD 10-OCT-2002.
XX PF 28-MAR-2002; 2002WO-US009671.
XX PR 28-MAR-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.
XX PR 01-OCT-2001; 2001US-0326370P.
XX PR 04-DEC-2001; 2001US-0336780P.
XX PR 20-FEB-2002; 2002US-0358985P.
XX PA (ZYCO-) ZYCOS INC.
XX PI Chicx RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX Example 2; SEQ ID NO 1823; 134pp; English.
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
```

```
SQ Sequence 445 AA;
Query Match      100.0%; Score 2071; DB 6; Length 445;
Best Local Similarity 100.0%; Pred. No. 8.5e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDFAMDVYKNLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFDPETTRQMOSTRD 60
Db 50 MYDFAMDVYKNLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFDPETTRQMOSTRD 109
QY 61 GRMLFDYLADKHGFRQYLDLTLYRYAKFOYECGNYSAAEYLYFFRVLPATDRNALSSL 120
Db 110 GRMLFDYLADKHGFRQYLDLTLYRYAKFOYECGNYSAAEYLYFFRVLPATDRNALSSL 169
QY 121 WCKLASEILMQNWDAMEDLTRLKETIDNNSVSSPLQSLQORTWLIHNSLFVFFNHPKGR 180
Db 170 WCKLASEILMQNWDAMEDLTRLKETIDNNSVSSPLQSLQORTWLIHNSLFVFFNHPKGR 229
QY 181 DNIIDFLYQOYLNAIQWCPHILRYLTAVITNKDVRKRQVLKDLVKVIQOESYTYK 240
Db 230 DNIIDFLYQOYLNAIQWCPHILRYLTAVITNKDVRKRQVLKDLVKVIQOESYTYK 289
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
Db 290 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 349
QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 350 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 409
QY 361 SLSFRSQMLAMNIEKKLNONSSEAPNWTQDSGFY 396
Db 410 SLSFRSQMLAMNIEKKLNONSSEAPNWTQDSGFY 445
```

Search completed: March 8, 2005, 20:04:18  
Job time : 167 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 20:17:10 ; Search time 174 Seconds

(without alignments)

1165.422 Million cell updates/sec

Title: US-10-783-415-4

Perfect score: 396

Sequence: 1 MVDPMVYKNLYSDIPHA.....LNQNSRSEAPNWTQDSGCFY 396

Scoring table: <BFGGO>

Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt 03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	396	100.0	396	Q9BRV2	Q9brv2 homo sapien
2	396	100.0	445	1 IF36 HUMAN	P60228 homo sapien
3	396	100.0	445	1 IF36 MOUSE	P60229 mus musculus
4	396	100.0	445	2 Q641X8	Q641x8 rattus norv
5	373	94.2	422	2 Q9CT23	Q9ct23 mus musculus
6	295	74.5	445	2 Q8WVK4	Q8wvk4 homo sapien
7	295	74.5	445	2 Q6TAX5	Q6tax5 homo sapien
8	143	36.1	423	2 Q7ZXA5	Q7zxa5 xenopus lae
9	143	36.1	446	2 Q6P7L9	Q6p7l9 xenopus tro
10	105	26.5	119	2 Q8BNE6	Q8bne6 mus musculus
11	79	19.9	446	2 Q918W7	Q918w7 xenopus lae
12	73	18.4	73	2 Q6LCK2	Q6lck2 mus musculus
13	72	18.2	446	2 Q6DR11	Q6dr11 brachydanio
14	72	18.2	446	2 Q6P7X8	Q6p7x8 brachydanio
15	72	18.2	448	2 Q6TH16	Q6th16 brachydanio
16	25	6.3	462	2 Q7Q1L8	Q7q1l8 anopheles g
17	15	3.8	418	2 Q9M2L8	Q9m2l8 arabidopsis
18	15	3.8	441	2 Q9C5Z3	Q9c5z3 arabidopsis
19	15	3.8	441	2 Q9M4T7	Q9m4t7 arabidopsis
20	13	3.3	372	2 Q8MR88	Q8mr88 drosophila
21	13	3.3	435	1 IF36 DROME	Q77410 drosophila
22	10	2.5	397	2 Q6W37	Q6w37 oryza sativ
23	10	2.5	415	2 Q69W36	Q69w36 oryza sativ
24	10	2.5	439	2 Q8GV10	Q8gv10 oryza sativ
25	9	2.3	421	2 Q8CQ02	Q8cq02 yarrowia li
26	9	2.3	517	2 Q81315	Q81315 plasmodium
27	9	2.3	527	2 Q7RQ22	Q7rq22 plasmodium
28	8	2.0	136	2 Q7LZ00	Q7lz00 ovine lenti
29	8	2.0	141	2 Q841L1	Q841l1 streptomyce
30	8	2.0	175	1 SRP PROMP	Q7uzp2 prochloroc
31	8	2.0	202	2 Q6DFE2	Q6dfe2 xenopus lae

32	8	2.0	215	2	Q6D0T5	Q6d0t5 erwinia car
33	8	2.0	299	2	Q6FB96	Q6fb96 acinetobact
34	8	2.0	335	1	COBT_TRETH	Q781c7 thermus the
35	8	2.0	336	2	Q746P7	Q746p7 thermus the
36	8	2.0	368	2	Q6Y686	Q6y686 mus musculu
37	8	2.0	393	2	Q9PLP1	Q9plp1 chlamydia m
38	8	2.0	395	2	Q6Y687	Q6y687 homo sapien
39	8	2.0	425	1	GSPL_ERWCA	P31708 erwinia car
40	8	2.0	425	2	Q6D2J4	Q6d2j4 erwinia car
41	8	2.0	432	1	IF36_CABEL	Q61820 caenorhabd
42	8	2.0	432	2	Q9Q8W7	Q9q8w7 rabbit fibr
43	8	2.0	436	2	Q6Z475	Q6z475 oryza sativ
44	8	2.0	436	2	Q9LNL2	Q9lnl2 arabidopsis
45	8	2.0	493	1	AMPA_CLOPE	Q8xh13 clostridium

#### ALIGNMENTS

#### RESULT 1

Q9BRV2 PRELIMINARY; PRT; 396 AA.  
AC Q9BRV2;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE EIF356 protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.N., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch J.E.,  
RT Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RA Strausberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC005944; AAH05944.1; -.  
DR InterPro; IPR001717; PCI.  
DR InterPro; IPR010935; SMC\_hinge.  
DR Pfam; PF01399; PCI; 1.  
DR SMART; SM00088; PINT; 1.  
SQ SEQUENCE 396 AA; 46445 MW; 09165D0D612B1C3B CRC64;

Query Match 100.0%; Score 396; DB 2; Length 396;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDPMVYKNLYSDIPHA...LNQNSRSEAPNWTQDSGCFY 396

|||||

Db 1 MVDFAVDYKLYNSDDIPHALREKRTTVAQLKQAEETEPVVKFDEPDTTROMQSTRD 60  
Qy 61 GRMLFDYLADKHGRQEQYDLTYAKFQYECGNSGAAEYLYFRVLVLPATDNLSSSL 120  
Db 61 GRMLFDYLADKHGRQEQYDLTYAKFQYECGNSGAAEYLYFRVLVLPATDNLSSSL 120  
Qy 121 WGLKASEILMQNDAAMEDLTLKKTETDNNSSVSPLOSLQORTWLHWSLVFVFNHPKGR 180  
Db 121 WGLKASEILMQNDAAMEDLTLKKTETDNNSSVSPLOSLQORTWLHWSLVFVFNHPKGR 180  
Qy 181 DNIIDLFLYQOYLNAITCMTCPHILRLYTTAVITNKDVKRQVQLKLVKVIQESYTYK 240  
Db 181 DNIIDLFLYQOYLNAITCMTCPHILRLYTTAVITNKDVKRQVQLKLVKVIQESYTYK 240  
Qy 241 DPITEFVECLVNFDFDGAQKLRCECVLVNDFLVACLEDFTENARLFIFETFCRIHQ 300  
Db 241 DPITEFVECLVNFDFDGAQKLRCECVLVNDFLVACLEDFTENARLFIFETFCRIHQ 300  
Qy 301 C1SINMLADKLNMTPEEAERIVNLIRNARDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360  
Db 301 C1SINMLADKLNMTPEEAERIVNLIRNARDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360  
Qy 361 SLSFRSOMLAWNIIEKLNQNSRSEAPNATQDSGFY 396  
Db 361 SLSFRSOMLAWNIIEKLNQNSRSEAPNATQDSGFY 396

## RESULT 2

IF36\_HUMAN STANDARD; PRT; 445 AA.  
ID P60228; O43902; Q64058; Q64059; Q64252;  
DC 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Eukaryotic translation initiation factor 3 subunit 6 (eIF-3 p48)  
DE (eIF3e) (Viral integration site protein INT-6 homolog).  
GN Name=EIF3S6; Synonyms=INT6;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 280-289 AND 427-436.  
RC TISSUE=Liver;  
RX MEDLINE=97442403; PubMed=9295280; DOI=10.1074/jbc.272.38.23477;  
RA Asano K., Merrick W.C., Hershey J.W.B.;  
RT "The translation initiation factor eIF3-p48 subunit is encoded by int-6, a site of frequent integration by the mouse mammary tumor virus genome.";  
RT J. Biol. Chem. 272:23477-23480 (1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=98066777; PubMed=9403073; DOI=10.1006/geno.1997.4996;  
RA Miyazaki S., Imatani A., Ballard L., Marchetti A., Buttitta F.,  
RA Albertsen H., Nevanlinna H.A., Gallahan D., Callahan R.;  
RT "The chromosome location of the human homolog of the mouse mammary tumor-associated gene INT6 and its status in human breast carcinomas.";  
RT Genomics 46:155-158 (1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96337990; PubMed=8688078;  
RA Desbois C., Rousset R., Bantignies F., Jalinot P.;  
RT "Exclusion of int-6 from PML nuclear bodies by binding to the HTLV-I Tax oncoprotein.";  
RT Science 273:951-953 (1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX PubMed=12386384;  
RA Neuvirt C., Jin D.-Y., Semmes O.J., Diella F., Callahan R.,  
RA Jeang K.-T.;  
RT "Divergent subcellular locations of HTLV-I Tax and Int-6: a contrast

between in vitro protein-protein binding and intracellular protein colocalization.";  
J. Biomed. Sci. 4:229-234 (1997).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow, Brain, and Muscle;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klusner R.F., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.N., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [6]  
RP INTERACTION WITH TRIM27.  
RX PubMed=10504338;  
RA Morris-Desbois C., Bochar V., Reynaud C., Jalinot P.;  
RT "Interaction between the Ret finger protein and the int-6 gene product and co-localisation into nuclear bodies.";  
RL J. Cell Sci. 112:3331-3342 (1999).  
CC -!- FUNCTION: Binds to the 40S ribosome and promotes the binding of methionyl-tRNAi and mRNA.  
CC -!- SUBUNIT: eIF-3 is composed of at least 12 different subunits. Interacts with TRIM27.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.  
CC -!- SIMILARITY: Belongs to the EIF3S6 family.  
CC -!- SIMILARITY: Contains 1 PCI domain.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; U54562; AAC51760.1; -  
EMBL; U94174; AAC51917.1; -  
EMBL; U94162; AAC51917.1; JOINED.  
EMBL; U94163; AAC51917.1; JOINED.  
EMBL; U94164; AAC51917.1; JOINED.  
EMBL; U94165; AAC51917.1; JOINED.  
EMBL; U94166; AAC51917.1; JOINED.  
EMBL; U94167; AAC51917.1; JOINED.  
EMBL; U94168; AAC51917.1; JOINED.  
EMBL; U94169; AAC51917.1; JOINED.  
EMBL; U94170; AAC51917.1; JOINED.  
EMBL; U94171; AAC51917.1; JOINED.  
EMBL; U94172; AAC51917.1; JOINED.  
EMBL; U94173; AAC51917.1; JOINED.  
EMBL; U94175; AAC51919.1; -  
EMBL; U62962; AAB58251.1; -  
EMBL; U85947; AAB88873.1; -  
EMBL; BC000734; AAH00734.1; -  
EMBL; BC008419; AAH08419.1; -  
EMBL; BC016706; AAH16706.1; -  
EMBL; BC021679; AAH21679.1; -  
IntAct; P60228; -

DR Genew; HGNC:3277; EIF3S6.  
 DR H-InvDB; HIX0007722; --  
 DR Reactome; P60228; --  
 DR MIM; 602210; --  
 DR InterPro; IPR000717; PCI.  
 KW Direct protein sequencing; Initiation factor; Protein biosynthesis.  
 SQ SEQUENCE 445 AA; 52220 MW; A5368651DDDDDD0C CRC64;

Query Match 100.0%; Score 396; DB 1; Length 445;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDYKLYSDDIIPALREKKTIVVAQLKQLQAEETPVKMFEDPETTRQMOSTRD 60  
 DB 50 MVDFAVDYKLYSDDIIPALREKKTIVVAQLKQLQAEETPVKMFEDPETTRQMOSTRD 109

QY 61 GRMLFDYLADKHGFRQYEDLTLYRYAKFYQECGNSGAAEYLYFFRVLVVPATDRNALSSL 120  
 DB 110 GRMLFDYLADKHGFRQYEDLTLYRYAKFYQECGNSGAAEYLYFFRVLVVPATDRNALSSL 169

QY 121 WGLKASEILMQNDAAEDTLRLKETIDNNSVSSPLOSLOQTWLIHWSLFFVFNHPKGR 180  
 DB 170 WGLKASEILMQNDAAEDTLRLKETIDNNSVSSPLOSLOQTWLIHWSLFFVFNHPKGR 229

QY 181 DNIIDFLYQPOLNAIOTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240  
 DB 230 DNIIDFLYQPOLNAIOTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 289

QY 241 DPTEFVCLYVDFDGAQKLRCESVLVNDFELVACLEDFIENARLFIPTFCRIHQ 300  
 DB 290 DPTEFVCLYVDFDGAQKLRCESVLVNDFELVACLEDFIENARLFIPTFCRIHQ 349

QY 301 CISINMLADKLMTPPEAEIRIVNLIRNARLDKIDSKLGHVVMGNNAVSPVQVIEKTK 360  
 DB 350 CISINMLADKLMTPPEAEIRIVNLIRNARLDKIDSKLGHVVMGNNAVSPVQVIEKTK 409

QY 361 SLSFRSOLAMNIEKLNQSRSEAPNATQDSGFY 396  
 DB 410 SLSFRSOLAMNIEKLNQSRSEAPNATQDSGFY 445

RESULT 3  
 IF36 MOUSE  
 ID IF36 MOUSE STANDARD; PRT: 445 AA.  
 AC P60229; Q43902; Q64058; Q64059; Q64252;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DE Eukaryotic translation initiation factor 3 subunit 6 (eIF-3 p48)  
 DE (eIF3e) (Mammary tumor-associated protein INT-6) (Viral integration site protein INT-6) (MMTV integration site 6);  
 GN Name=EIF3e6; Synonyms=Int6;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95156630; PubMed=7853537;  
 RA Marchetti A., Buttitta F., Miyazaki S., Gallahan D., Smith G.H., Callahan R.;  
 RT "Int-6, a highly conserved, widely expressed gene, is mutated by mouse mammary tumor virus in mammary preneoplasia.";  
 RL J. Virol. 69:1932-1938(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97405883; PubMed=9260927;  
 RA Diella F., Levi G., Callahan R.;  
 RT "Characterization of the INT6 mammary tumor gene product.";  
 RL DNA Cell Biol. 16:839-847(1997).  
 RN [3]  
 RP REVISIONS TO N-TERMINUS.  
 RA Callahan R.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Czech II; TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Ussin T.B., Toshiyuki S., Carninci P., Prange C., Roach S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Baska S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Griewood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Binds to the 40S ribosome and promotes the binding of methionyl-tRNAi and mRNA.  
 CC -!- SUBUNIT: eIF-3 is composed of at least 12 different subunits. Interacts with TRIM27 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.  
 CC -!- DISEASE: Int-6 serves as a site for viral integration of mouse mammary tumor virus (MMTV) in mammary tumors.  
 CC -!- SIMILARITY: Belongs to the EIF36 family.  
 CC -!- SIMILARITY: Contains 1 PCI domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).  
 -----  
 CC EMBL; S75221; AAC00046.1; --  
 CC EMBL; S75223; AAC00047.1; --  
 DR EMBL; BC029177; AAH29177.1; --  
 DR MGI; MGI:99257; Eif3e6.  
 DR InterPro; IPR000717; PCI  
 KW Initiation factor; Protein biosynthesis; Proto-oncogene.  
 FT VARIANT 158 158  
 FT VARIANT 158 180  
 FT VARIANT 159 445  
 FT VARIANT 318 365  
 FT VARIANT 366 445  
 FT VARIANT 445 AA; 52220 MW; A5368651DDDDDD0C CRC64;  
 SQ SEQUENCE 445 AA; 52220 MW; A5368651DDDDDD0C CRC64;

Query Match 100.0%; Score 396; DB 1; Length 445;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDYKLYSDDIIPALREKKTIVVAQLKQLQAEETPVKMFEDPETTRQMOSTRD 60  
 DB 50 MVDFAVDYKLYSDDIIPALREKKTIVVAQLKQLQAEETPVKMFEDPETTRQMOSTRD 109

QY 61 GRMLFDYLADKHGFRQYEDLTLYRYAKFYQECGNSGAAEYLYFFRVLVVPATDRNALSSL 120  
 DB 110 GRMLFDYLADKHGFRQYEDLTLYRYAKFYQECGNSGAAEYLYFFRVLVVPATDRNALSSL 169

QY 121 WGLKASEILMQNDAAEDTLRLKETIDNNSVSSPLOSLOQTWLIHWSLFFVFNHPKGR 180  
 DB 170 WGLKASEILMQNDAAEDTLRLKETIDNNSVSSPLOSLOQTWLIHWSLFFVFNHPKGR 229

QY 181 DNIIDFLYQPOLNAIOTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240  
 DB 230 DNIIDFLYQPOLNAIOTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 289

QY 241 DPTEFVCLYVDFDGAQKLRCESVLVNDFELVACLEDFIENARLFIPTFCRIHQ 300  
 DB 290 DPTEFVCLYVDFDGAQKLRCESVLVNDFELVACLEDFIENARLFIPTFCRIHQ 349

QY 301 CISINMLADKLMTPPEAEIRIVNLIRNARLDKIDSKLGHVVMGNNAVSPVQVIEKTK 360  
 DB 350 CISINMLADKLMTPPEAEIRIVNLIRNARLDKIDSKLGHVVMGNNAVSPVQVIEKTK 409

QY 361 SLSFRSOLAMNIEKLNQSRSEAPNATQDSGFY 396  
 DB 410 SLSFRSOLAMNIEKLNQSRSEAPNATQDSGFY 445

```

Qy 121 WGLASEILMNQWDAAMEDLTRLKETIDNNSVSPQSQLOQRTWLIIHWSLFFVFNHPKGR 180
|
|
|
Db 170 WGLASEILMNQWDAAMEDLTRLKETIDNNSVSPQSQLOQRTWLIIHWSLFFVFNHPKGR 229
|
|
|
Qy 181 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240
|
|
|
Db 230 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 289
|
|
|
Qy 241 DPITEFVECLVNFDFDGAQKKLRECESVLVNDFLVLACLEDFIENARLFIETFCRIHQ 300
|
|
|
Db 290 DPITEFVECLVNFDFDGAQKKLRECESVLVNDFLVLACLEDFIENARLFIETFCRIHQ 349
|
|
|
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQVIEKTK 360
|
|
|
Db 350 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQVIEKTK 409
|
|
|
Qy 361 SLSPRSQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
|
|
|
Db 410 SLSPRSQMLANNIEKKLNQNSRSEAPNWTQDSGFY 445
|
|
|
RESULT 4
Q641X8 PRELIMINARY; PRT; 445 AA.
AC Q641X8;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein. (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RP [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Director MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC082087; AA082087.1; -.
KW Hypothetical protein.
SQ SEQUENCE 445 AA; 52220 MW; A5368651DD0DDDD0C CRC64;
Query Match 100.0%; Score 396; DB 2; Length 445;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVDFAVDYKYNLYSDDIPHALREKRTTVAQKQLQAEITEPIVKMFDEPTTRQMQSTRD 60
|
|
|
Db 50 MVDFAVDYKYNLYSDDIPHALREKRTTVAQKQLQAEITEPIVKMFDEPTTRQMQSTRD 109
|
|
|

```

```

Qy 61 GRMLFDYLADKHGFRQEVLDLTRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
|
|
|
Db 110 GRMLFDYLADKHGFRQEVLDLTRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 169
|
|
|
Qy 121 WGLASEILMNQWDAAMEDLTRLKETIDNNSVSPQSQLOQRTWLIIHWSLFFVFNHPKGR 180
|
|
|
Db 170 WGLASEILMNQWDAAMEDLTRLKETIDNNSVSPQSQLOQRTWLIIHWSLFFVFNHPKGR 229
|
|
|
Qy 181 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240
|
|
|
Db 230 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 289
|
|
|
Qy 241 DPITEFVECLVNFDFDGAQKKLRECESVLVNDFLVLACLEDFIENARLFIETFCRIHQ 300
|
|
|
Db 290 DPITEFVECLVNFDFDGAQKKLRECESVLVNDFLVLACLEDFIENARLFIETFCRIHQ 349
|
|
|
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQVIEKTK 360
|
|
|
Db 350 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQVIEKTK 409
|
|
|
Qy 361 SLSPRSQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
|
|
|
Db 410 SLSPRSQMLANNIEKKLNQNSRSEAPNWTQDSGFY 445
|
|
|
RESULT 5
Q9CT23 PRELIMINARY; PRT; 422 AA.
AC Q9CT23;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:2610017H09 product:mammary tumor integration
DE site 6, full insert sequence. (Fragment).
GN Name=B1f386;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=93279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RN Meth. Enzymol. 303:19-44 (1999).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/350555500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690 (2001).
RP [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RN Nature 420:563-573 (2002).
RP [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RN Genome Res. 10:1617-1630 (2000).
RP [5]

```

RC SEQUENCE FROM N.A.  
 RX STRAIN=C57BL/6J; TISSUE=Whole body;  
 RA MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
 RA "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Akawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK011436; BAB27621.1; -;  
 DR MGD; MGI:99257; Eif3a6.  
 DR InterPro; IPR000717; PCI.  
 DR InterPro; IPR010935; SMC\_hinge.  
 DR Pfam; PF01399; PCI; 1.  
 DR SMART; SMO0089; PINT; 1.  
 FT NON TER 422 422  
 SQ SEQUENCE 422 AA; 49568 MW; A05157DBCC6D375B CRC64;

Query Match 94.2%; Score 373; DB 2; Length 422;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVDFAVDYKNLYSDDI PHALREKRTTVAOLKQLOAETEPVVKMFDPETTRQMOSTRD 60  
 DB 50 MVDFAVDYKNLYSDDI PHALREKRTTVAOLKQLOAETEPVVKMFDPETTRQMOSTRD 109  
 QY 61 GRMLFDYLDKXGFRQYEDLTLYRYAKFYQYECGNYSGAAEYLYFFRVLVPAFDRNALSSL 120  
 DB 110 GRMLFDYLDKXGFRQYEDLTLYRYAKFYQYECGNYSGAAEYLYFFRVLVPAFDRNALSSL 169  
 QY 121 WGLKASEILMNQWDAAMEDLTRELKETIDNNSSVSSPLQSLQORTWLIIHWSLFFVFNHPKGR 180  
 DB 170 WGLKASEILMNQWDAAMEDLTRELKETIDNNSSVSSPLQSLQORTWLIIHWSLFFVFNHPKGR 229  
 QY 181 DNIIDLFLYQYQLNAIOTMCPHILRYLTAVITNKDKVRKROVLKDLVKVIOQESYTYK 240  
 DB 230 DNIIDLFLYQYQLNAIOTMCPHILRYLTAVITNKDKVRKROVLKDLVKVIOQESYTYK 289  
 QY 241 DPITEFVCLYVNFDFDGAQKLRCEESVNVNDFLVACLEDFIENARLFIETFCRIHQ 300  
 DB 290 DPITEFVCLYVNFDFDGAQKLRCEESVNVNDFLVACLEDFIENARLFIETFCRIHQ 349  
 QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
 DB 350 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 409  
 QY 361 SLFSRQMLAMNI 373  
 DB 410 SLFSRQMLAMNI 422

RESULT 6

Q8WVK4

ID Q8WVK4 PRELIMINARY; PRT; 445 AA.

Q8WVK4;  
 AC 01-WAR-2002 (TREMBlrel. 20, Created)  
 DT 01-WAR-2002 (TREMBlrel. 20, Last sequence update)  
 DE 01-WAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Mammary tumor integration site 6 (Oncogene homolog).  
 GN Name=Eif3a6;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Pridge C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC017887; AAH17887.1; -;  
 DR Pfam; PF01399; PCI; 1.  
 DR SMART; SMO0089; PINT; 1.  
 SQ SEQUENCE 445 AA; 52248 MW; AFOAD651DAAA8D0C CRC64;  
 Query Match 74.5%; Score 295; DB 2; Length 445;  
 Best Local Similarity 99.7%; Pred. No. 1.4e-299;  
 Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MVDFAVDYKNLYSDDI PHALREKRTTVAOLKQLOAETEPVVKMFDPETTRQMOSTRD 60  
 DB 50 MVDFAVDYKNLYSDDI PHALREKRTTVAOLKQLOAETEPVVKMFDPETTRQMOSTRD 109  
 QY 61 GRMLFDYLDKXGFRQYEDLTLYRYAKFYQYECGNYSGAAEYLYFFRVLVPAFDRNALSSL 120  
 DB 110 GRMLFDYLDKXGFRQYEDLTLYRYAKFYQYECGNYSGAAEYLYFFRVLVPAFDRNALSSL 169  
 QY 121 WGLKASEILMNQWDAAMEDLTRELKETIDNNSSVSSPLQSLQORTWLIIHWSLFFVFNHPKGR 180  
 DB 170 WGLKASEILMNQWDAAMEDLTRELKETIDNNSSVSSPLQSLQORTWLIIHWSLFFVFNHPKGR 229  
 QY 181 DNIIDLFLYQYQLNAIOTMCPHILRYLTAVITNKDKVRKROVLKDLVKVIOQESYTYK 240  
 DB 230 DNIIDLFLYQYQLNAIOTMCPHILRYLTAVITNKDKVRKROVLKDLVKVIOQESYTYK 289  
 QY 241 DPITEFVCLYVNFDFDGAQKLRCEESVNVNDFLVACLEDFIENARLFIETFCRIHQ 300  
 DB 290 DPITEFVCLYVNFDFDGAQKLRCEESVNVNDFLVACLEDFIENARLFIETFCRIHQ 349  
 QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360  
 DB 350 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 409  
 QY 361 SLFSRQMLAMNI 422

```
|||||
410 SLSPRSQMLAMNIEKLNQSRSEAPNWTQDSGFY 445

RESULT 7
Q61AX5
ID Q61AX5 PRELIMINARY; PRT; 445 AA.
AC Q61AX5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE EIF356 protein.
GN Name=EIF356;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OK NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR457029; CAG33310.1; -.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
SQ SEQUENCE 445 AA; 52205 MW; C7C0PF755BD9A8D4 CRC64;

Query Match 74.5%; Score 295; DB 2; Length 445;
Best Local Similarity 99.7%; Pred. No. 1.4e-299;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMQSTRD 60
Db 50 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMQSTRD 109
Qy 61 GRMLFDYLADKHGRFQRYEYLDLYYAKFQYECGNYSGAAEYLYFPRVLVPATDRLNLSL 120
Db 110 GRMLFDYLADKHGRFQRYEYLDLYYAKFQYECGNYSGAAEYLYFPRVLVPATDRLNLSL 169
Qy 121 WGLKASEILMNQWDAAMEDLRLKETIDNNSSVSPQSLQSLQRTWLIIHWSLVFFNHPKGR 180
Db 170 WGLKASEILMNQWDAAMEDLRLKETIDNNSSVSPQSLQSLQRTWLIIHWSLVFFNHPKGR 229
Qy 181 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKQVLDKLVKVIQESYTYK 240
Db 230 DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKQVLDKLVKVIQESYTYK 289
Qy 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLEDFIENARLFIFETFCRIHQ 300
Db 290 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLEDFIENARLFIFETFCRIHQ 349
Qy 301 CISINMLADKLNMTPEEAERIVNLIRNARLDKIDS KLGHVVMGNNAVSPYQVIEKTK 360
Db 350 CISINMLADKLNMTPEEAERIVNLIRNARLDKIDS KLGHVVMGNNAVSPYQVIEKTK 409
Qy 361 SLSPRSQMLAMNIEKLNQSRSEAPNWTQDSGFY 396
Db 410 SLSPRSQMLAMNIEKLNQSRSEAPNWTQDSGFY 445

RESULT 8
Q7ZXAS
ID Q7ZXAS PRELIMINARY; PRT; 423 AA.
AC Q7ZXAS;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE LOC398503 protein (Fragment).
GN Name=LOC398503;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

```

```
Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raba S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045079; AAH45079.1; -.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
FT NON TER 1
SQ SEQUENCE 423 AA; 49523 MW; F691314F43EF400B CRC64;

Query Match 36.1%; Score 143; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 1.9e-140;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 162 RTWLIIHWSLVFFNHPKGRDNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKR 221
Db 189 RTWLIIHWSLVFFNHPKGRDNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKR 248
Qy 222 ROVLAKDLVKVILQOESYTYKDPITFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLE 281
Db 249 ROVLAKDLVKVILQOESYTYKDPITFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLE 308
Qy 282 DFIENARLFIFETFCRIHQICSI 304
Db 309 DFIENARLFIFETFCRIHQICSI 331

RESULT 9
Q6P7L9
ID Q6P7L9 PRELIMINARY; PRT; 446 AA.
AC Q6P7L9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Eukaryotic translation initiation factor 3, subunit 6..

```

GN Name=elf3-p48-prov;  
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8364;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=2238857; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altchek S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.P., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Gerhard D.S.;  
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC061611; AAH61611.1;  
DR GO: GO:0003743; P:translational initiation factor activity; IEA.  
DR InterPro; IPR000717; PCI.  
DR InterPro; IPR010935; SMC\_hinge.  
DR Pfam; PF01399; PCI; 1.  
DR SMART; SM00088; PINT; 1.  
KW Initiation factor.  
SQ SEQUENCE 446 AA; 52296 MW; EB12CE70888DA8B1 CRC64;  
Query Match 36.1%; Score 143; DB 2; Length 446;  
Best Local Similarity 100.0%; Pred. No. 2e-140;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 162 RTWLTHWSLFVFNHPKGRDNIIDLFLYQPVYLNAIQTMCPHLRYLTAVITNKDVRKR 221  
DB 212 RTWLTHWSLFVFNHPKGRDNIIDLFLYQPVYLNAIQTMCPHLRYLTAVITNKDVRKR 271  
QY 222 RQVLKDLVKVIOQESYTYKDPITEFVECLVYNFDFDGAQKLRCESVLVNDFFLVACLE 281  
DB 272 RQVLKDLVKVIOQESYTYKDPITEFVECLVYNFDFDGAQKLRCESVLVNDFFLVACLE 331  
QY 282 DFENARLFIFFTFCRIHQICSI 304  
DB 332 DFENARLFIFFTFCRIHQICSI 354  
RESULT 10  
QBNE6  
ID QBNE6 PRELIMINARY; PRT; 119 AA.  
AC QBNE6;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length  
DE enriched library, clone:D130036L03 product:mammary tumor integration  
DE site 6, full insert sequence.  
GN Name=Elf3e;  
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=C57BL/6J; TISSUE=Spinal ganglion;  
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
RA The FANTOM Consortium;  
RT the RIKEN Genome Exploration Research Group Phase I & II Team;  
RL "Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs.";  
RN Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
RX MEDLINE=20530913; PubMed=11078661; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Alizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
RA Adachi J., Alizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
RA Horii F., Imotani K., Ishii Y., Ito H., Kagawa I., Kasukawa T.,  
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akai H., Tanaka Y., Tanaka T.,  
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK083874; BAC39045.1;  
DR MGI; MGI:99257; Bf386.  
DR InterPro; IPR010935; SMC\_hinge.  
SQ SEQUENCE 119 AA; 14113 MW; ACC86F6ABBE052E6 CRC64;  
Query Match 26.5%; Score 105; DB 2; Length 119;  
Best Local Similarity 100.0%; Pred. No. 3.5e-101;  
Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 130 MQNDAMEDLTRKETIDNNSSVSLQSLQRTWLHWSLFVFNHPKGRDNIIDLFLY 189

Db 1 MONWDAAMEDTLRLKETIDNNSVSPLOSQORTWLIHWSLFVFNHFKGRDNIIDFLY 60  
Qy 190 QPQYLNAIQTWCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQ 234  
Db 61 QPQYLNAIQTWCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQ 105

## RESULT 11

Q918W7 PRELIMINARY; PRT; 446 AA.  
AC Q918W7;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
DE Int-6 protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Morris-Desbois C., Bochar V., Reynaud C., Jalinot P.;  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF162775; AAF80474.1; -.  
DR InterPro; IPR000717; PCL\_hinge.  
DR InterPro; IPR010935; SMC\_hinge.  
DR Pfam; PF01399; PCI; 1.  
DR SMART; SMO0088; PINT; 1.  
SQ SEQUENCE 446 AA; 52240 MW; 21C5605DA9CF88BF CRC64;  
Query Match 19.9%; Score 79; DB 2; Length 446;  
Best Local Similarity 100.0%; Pred. No. 2e-73;  
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 181 DNIIDFLYQOYLNAIQTWCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240  
Db 231 DNIIDFLYQOYLNAIQTWCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 290  
Qy 241 DPITEFVECLYVNFDFGA 259  
Db 291 DPITEFVECLYVNFDFGA 309

## RESULT 12

Q6LCK2 PRELIMINARY; PRT; 73 AA.  
AC Q6LCK2;  
DT 05-JUL-2004 (TReMBLrel. 27, Created)  
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
DE Int-6 protein (Fragment).  
GN Name=Int-6;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97442403; PubMed=9295280; DOI=10.1074/jbc.272.38.23477;  
RA Asano K., Merrick W.C., Hershey J.W.;  
RT "The translation initiation factor eIF3-p48 subunit is encoded by int-6, a site of frequent integration by the mouse mammary tumor virus genome."  
RT J. Biol. Chem. 272:23477-23480 (1997).  
RL EMBL; U54563; AAC53346.1; -.  
FT NON TER 73  
SQ SEQUENCE 73 AA; 8565 MW; 0904AD1E79842124 CRC64;

Query Match 18.4%; Score 73; DB 2; Length 73;  
Best Local Similarity 100.0%; Pred. No. 7.2e-68;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDYKXLYSDDIPHALREKRTTVAQKQLQAEETPIVKMFEDPETTRQMQSTRD 60  
Db 1 MVDFAVDYKXLYSDDIPHALREKRTTVAQKQLQAEETPIVKMFEDPETTRQMQSTRD 60  
Qy 61 GRMLFDYLDADKHG 73  
Db 61 GRMLFDYLDADKHG 73

## RESULT 13

Q6DRII PRELIMINARY; PRT; 446 AA.  
AC Q6DRII;  
DT 25-OCT-2004 (TReMBLrel. 28, Created)  
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)  
DE Eukaryotic translation initiation factor 3 subunit 6.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA PubMed=15256591; DOI=10.1073/pnas.0403929101;  
RX Amsterdams A., Nissen R.M., Sun Z., Swindell E.C., Farrington S.,  
RA Hopkins N.;  
RT "Identification of 315 genes essential for early zebrafish development."  
RL EMBL; AY648778; AAT68096.1; -.  
DR GO; GO:0003743; F-translation initiation factor activity; IEA.  
DR InterPro; IPR000717; PCL\_hinge.  
DR InterPro; IPR010935; SMC\_hinge.  
DR Pfam; PF01399; PCI; 1.  
DR SMART; SMO0088; PINT; 1.  
KW Initiation factor.  
SQ SEQUENCE 446 AA; 52398 MW; 2C36E460A8002BB7 CRC64;

Query Match 18.2%; Score 72; DB 2; Length 446;  
Best Local Similarity 100.0%; Pred. No. 4.3e-66;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 FEDPETTRQMQSTRDGRMLFDYLDADKHGFRQYLDTRYAKFYECGNYSGAAYLYFF 105  
Db 96 FEDPETTRQMQSTRDGRMLFDYLDADKHGFRQYLDTRYAKFYECGNYSGAAYLYFF 155  
Qy 106 RVLVPATDRNAL 117  
Db 156 RVLVPATDRNAL 167

## RESULT 14

Q6P7X8 PRELIMINARY; PRT; 446 AA.  
AC Q6P7X8;  
DT 05-JUL-2004 (TReMBLrel. 27, Created)  
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
DE Eukaryotic translation initiation factor 3, subunit 6.  
GN ORFNames=zgc:63821;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB; TISSUE=Whole body;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,



RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB; TISSUE=whole body;  
RA Strausberg R.;  
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC061454; AAH61454.1; -.  
DR ZFIN; ZDB-GENE-030131-3827; zgc:63821.  
DR GO; GO:0003743; P:translation initiation factor activity; IEA.  
DR InterPro; IPR000717; PCI.  
DR InterPro; IPR010935; SMC\_hinge.  
DR Pfam; PF01399; PCI; 1.  
DR SMART; SM00088; PINT; 1.  
KW Initiation factor.  
SQ SEQUENCE 446 AA; 52389 MW; 32F9D000E69F4F87 CRC64;

Query Match 18.2%; Score 72; DB 2; Length 446;  
Best Local Similarity 100.0%; Pred. No. 4.3e-66;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 46 FEDPTTQMOSTRQGRMLFDYLDKQKGRQRYLDTLYRYAKFOYECGNYSGAAEYLYFF 105  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 96 FEDPTTQMOSTRQGRMLFDYLDKQKGRQRYLDTLYRYAKFOYECGNYSGAAEYLYFF 155  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
  
QY 106 RVLVPATDRNAL 117  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 156 RVLVPATDRNAL 167

RESULT 15  
Q6TH16  
ID Q6TH16 PRELIMINARY; PRT; 448 AA.  
AC Q6TH16;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Eukaryotic translation initiation factor 3, subunit 6 48kDa.  
GN Name=EIF3S6; ORFNames=zgc:63821;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney marrow;  
RA Song H.D., Wu X.Y., Sun X.J., Zhou Y., Liu T.X., Deng M., Zhang G.W.,  
RA Sheng Y., Chen Y., Ruan Z., Jiang C.L., Fan H.Y., Zou L.I.,  
RA Kanki J.P., Look A.T., Chen Z.;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY398340; AAQ97773.1; -.  
DR ZFIN; ZDB-GENE-030131-3827; zgc:63821.  
DR GO; GO:0003743; P:translation initiation factor activity; IEA.  
DR InterPro; IPR000717; PCI.  
DR InterPro; IPR010935; SMC\_hinge.  
DR Pfam; PF01399; PCI; 1.

DR SMART; SM00088; PINT; 1.  
KW Initiation factor.  
SQ SEQUENCE 448 AA; 52672 MW; 190C9FD4CEB77803 CRC64;  
  
Query Match 18.2%; Score 72; DB 2; Length 448;  
Best Local Similarity 100.0%; Pred. No. 4.3e-66;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 186 LFLYQFQYLNAIQTMCPHILRYLTAVITNKQVRKRQVLKDLVKVIOQESYTYKDPITE 245  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 238 LFLYQFQYLNAIQTMCPHILRYLTAVITNKQVRKRQVLKDLVKVIOQESYTYKDPITE 297  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
  
QY 246 FVECLYVNFDFD 257  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 298 FVECLYVNFDFD 309

Search completed: March 8, 2005, 20:26:35  
Job time : 177 secs

**THIS PAGE BLANK (USPTO)**